

Issue 8(1) Autumn Conference 2014

(Print) ISSN 1748-9393
(Online) ISSN 2051-5379

*British
Journal of
Undergraduate
Philosophy*



Editor-in-chief: Daniel Bregman
University of Oxford

Journal of the British Undergraduate Philosophy Society

British Journal of Undergraduate Philosophy

The Journal of the British Undergraduate Philosophy Society

Issue 8(1) Autumn 2014

Editor-in-chief:

Daniel Bregman

Editors:

Geoff Keeling

Dan Treger

Farbod Akhlaghi-Ghaffarokh

Manuscript editors:

Bede Hager-Suart

Nathan Oseroff

Sophie Osiecki

Andy Bates

Christopher Cheung

Jamie Parr

President of BUPS:

Dan Treger

Assistant President of BUPS: Farbod Akhlaghi-Ghaffarokh

www.bups.org/bjup-online

(Print) ISSN 1748-9393 | (Online) ISSN 2051-5359

British Journal of Undergraduate Philosophy

www.bups.org/bjup-online

All rights reserved.

© 2014 British Undergraduate Philosophy Society, and the authors.

No part of the *British Journal of Undergraduate Philosophy* may be reproduced or used in any form or by any means (graphic, electronic, or mechanical, including photocopying, taping, recording or by any other information storage and retrieval system) without written permission from the British Undergraduate Philosophy Society, and the authors.

(Print) ISSN 1748-9393 | (Online) ISSN 2051-5359

Editorial

What is bad for the bee, cannot be good for the hive
— Marcus Aurelius

Putting together a journal has oftentimes lead to comparisons between us and Socrates's image of the philosopher as midwife. We act as midwives, in some sense, delivering the work of others to the world. The analogy, of course, isn't perfect because the authors' pain in creating their work must surely pale in the face of ours in actually organising a journal, but, six months after our last edition, we stand ready with what you see before you.

We wanted to take this time to make a joint statement about the past and the future of the British Undergraduate Society and the *British Journal of Undergraduate Philosophy*, whose fates are inextricably bound together. The last few months have seen many bold changes and the introduction of new ideas and blood into the both the society and the journal team. But, before we plunge further into the unknown future, we want to take this moment to reflect upon the past, since those who neglect history are doomed to repeat it.

The society and journal have endured, though not necessarily perdured, through many leaders and styles over the years. What the last six months has seen is, primarily, an effort to stabilise this process and make it immune to the comings and goings of personnel that is so typical in an undergraduate society. In order to do so, positions have been clearly demarcated and a procedure of training assistants to take over a role has been effectively put in place. On the journal side, we have seen a streamlining of the processes that go into organising reviewers and quickening the turnaround we get on submissions. We are so pleased that after the riotous push of last minute details and worries that culminated in our last conference at Sheffield (which was still, no doubt, one of our better conferences), we were able to stretch out and calmly approach much of the process that has led to the current edition of the journal and the conference at which it is released. Really, there can be no greater sign of forward progress than this.

Every bit of this fantastic achievement is down to the hard work of people who get no reward for their work other than the joy of the finished product and the stoking of that philosophical fire that burns within us all. Specifically, we want to mention the inimitable Geoff Keeling who dived into the challenge of manuscript editing like a nubile fish to water. His impeccable organisational skills and cheery demeanour have made a sometimes nightmarish process one of genuine wonder. We thank him for his diligent work, and we look forward to his long and productive career with the Journal! A special mention must

also go out to Hannah Foulstone for heading up an entirely new branch of the society and forging it into something that will hopefully last far into the future as a necessary appendage of our societal outreach. Last but, but by no means least, we must give a deep, heartfelt, thanks to Nathan Oseroff who had the unenviable task of compiling the words arrayed before you in a very user friendly program known as L^AT_EX. We stand immensely indebted to his superior knowledge and skill and are only sad that we have mere words with which to thank him!

As always, this period marks the beginnings of many new careers and the ending of others. Alongside those bright sparks we mentioned above are others who are approaching the end of their life with the BUPS. One such person is Aliénor Cros, whom we are deeply saddened to see depart. Although she was thrust into her role of conference coordinating rather suddenly, she also outperformed herself and did incredible work in helping to put together three wonderful conferences. Aliénor, we are, to you, most grateful.

We look forward to the future of the British Undergraduate Philosophy Society with great enthusiasm. To be at the helm of an ever-changing and dispersed group as the Committee of this Society is both a challenge and a joy. We hope that over the coming term of presidency the Society will take strides to build upon the structural foundations that have been laid during the previous president's tenure. With such a structure, and with a watchful eye and centralising of communication, we are confident that the Society will thrive and fulfil one of its constitutional roles of providing a platform for undergraduates across the UK to meet, network, and listen to some of the finest of their peers present their work. It is, undoubtedly, a great honour and a privilege to be a part of a society such as this.

This year, we have experienced a significant increase of interest by submitting authors. Our intake of papers literally doubled to around 100 and a lot of the summer was spent organising and reviewing those papers. No matter how many editions of the journal you are privy to, you are always impressed by the great wealth of talent and ideas that lie out there in the vast world of UK Philosophy departments, and this year was no exception, producing some of the most interesting and well-thought-out essays we have seen from undergraduates. This year, the majority of accepted papers skewed towards the ethical side of things, and within that it was very interesting to read everything from meta-ethical considerations, to attempts to explain rights for animals as a basis for ethical behaviour, to questions about the moral status of objectification in pornography. Be the papers in the present edition ethics-related or not, we are confident that you will be impressed with their quality, as were we!

We would like to take this brief opportunity to also thank all those that made the conference possible. Specifically, Hallvard Lillehammer and Nils Kurbis of Birkbeck College for their exceptional generosity and patience during the planning the conference. We must also thank Baroness O'Neill and MM McCabe for taking time out of their busy schedules to speak to us. Personally, Daniel would like to thank his mother, Adele, for her support. He would also like to thank the department at KCL for their own, equally valued support. He would like to thank Ankita Anirban for sharing in his madness, Lydia Drabkin-Reiter for providing a sympathetic shoulder to lean on and especially Madeline Boden for always being there. Special thanks must also go to Max, Robin and Freddie for indulging in his vices with him. Farbod would like to thank his sister, Parisa, for her strength and care, and his parents, Forough and Fariborz, for their unwavering support of his pursuit of philosophy. He would also like to thank the Philosophy Department at the University of Reading for their support and generosity over the years. Finally, he would like to thank Adam for his friendship, for which he is most grateful, and Emma, for her support, on which he relies more than he readily admits.

Our reasons for doing what we do can sometimes seem mysterious to us. In moments of great pressure you might find yourself wondering why you put your name forward to take on the responsibility you bear. Indeed, at times, it might almost seem easy to shirk said responsibilities; a consideration of what binds you to your role may turn up little other than socially enforced obligation. But really what we do must partly stem from something much deeper than that. Those who find ourselves on this society and journal may do so for various reasons but we must firmly believe that we do what we do because of a fundamental appreciation or deep seated interest, perhaps even love, of philosophical expression; of the crystal clear presentation of genuinely intriguing ideas. We may talk about philosophical expression as a mode of human expression and it's all too easy to fall into the habit of thinking about 'philosophical' and 'proper' writing as two distinct categories of said expression. But reading the papers for this journal, one sees time and time again how we are all trying to capture similar ideas and themes – how, really, these are no more than two facets of how we look at the world and ourselves. Properly understood, philosophical expression can attain a beauty that one rarely finds outside of great novels. Its pursuit is certainly a most worthwhile task!

We sincerely hope you enjoy this issue of the *British Journal of Undergraduate Philosophy*.

DT, FA

Acknowledgements

A special thanks must be extended to our host of reviewers:

Alex Lloyd
Amna Whiston
Angelo Cei
Anna Ciborowska
Anna Ortin
Antoine Azzopardi
Chandler Clark
Christopher Burr
Christopher Ivins
Dan Woolley
David Allen
David Olbrich
Douglas Giles
Ekin Genc
Fotini Pantelides
Hannah Peacemen
Harry Lesser
Ian Robertson
Iancu Daramus
Jake Wojtowicz
James Andow
James Wenban
Jamie Parr
Jasmin Contos
Ji Yong Lee
Jinyu Sun
Jonathan Head
Jordan Rodger
Julia Tarnawska

Kai Kualulauu
Kevin Wallbridge
Lam Kei Yu Astrid
Lana Bracewell
Lenhardt Stevens
Leonidas Oikonomou
Marc Cole
Michael Bricker
Michael Plant
Mihnea Chiujea
Milo Nanni
Nataliya Bright
Nathan Oseroff
Nigel Pleasants
Noemi Magnani
Pete Faulconbridge
Peter West
Ravi Thakral
Rick Sendelbeck
Ryan Nefdt
Simon Young
Solveig Aasen
Sommer Rikke
Steven French
Suki Finn
Thomas Lee
Tom Schooen
Tom Winfield

Contents

- iii. Editorial
- vi. Acknowledgements
- 1. **Subsistence as a Basic Human Right:
A defence of Henry Shue**
Sean Butler
- 12. **A Set of Objections to David Benatar's
Anti-Natalism**
Franco Palazzi
- 37. **How Convincing is the Levelling Down Objec-
tion to Egalitarianism?**
Joel Griffith
- 47. **Killing the cat with too much cream?
A review of Donaldson and Kymlicka's
Zoopolis (OUP, 2013, 329 pp.)**
Collis Tahzib
- 57. ***The Dialectic of the Formal and the Intuitive*: Kurt
Gödel's Proof of the Ideality of Time**
Lukas Clark-Memler

67. **Michael Smith's 'Moral Problem' and a Defence of Metaethical Externalism**
Laura Nicoara

80. **On the Moral Status of Objectification in Pornography**
Carl-Otto Frietsch

88. **Can We Practically Measure Freedom?**
Alexander Tolcher

99. **Interview: Mary Margaret McCabe**

Subsistence as a Basic Human Right: A defence of Henry Shue

Sean Butler

University of St Andrews

In his *Basic Rights*, Henry Shue attempts to establish that: first, if there are any rights, there are rights to subsistence and physical security, because these are necessary for enjoying any other right; and, second, that all (basic) rights impose both positive and negative duties, and the two kinds of duty cannot be distinguished along these lines. On this basis, he argues that, if the global indigent are said to have any rights, they must therefore have a right to subsistence, and that this right imposes underived positive duties of aid on the global rich. Shue's first claim is based on his definition of a moral right, and the assertion that basic rights follow by conceptual necessity from this definition. This essay aims to rebut criticism of this first claim and argues that – on a suitable construal of Shue's definition that pays sufficient attention to (a) the normative element of his definition of a right and (b) construes 'enjoying a right' in a plausible way – Shue succeeds in showing that basic rights are entailed by non-basic rights, and that subsistence is one of them. The first section explains Shue's argument for basic rights; §2 examines Thomas Pogge's claim that there are no rights that could count as basic on any interpretation of 'enjoying a right' and argues that, while Pogge highlights a crucial ambiguity in what Shue means by this, there is a disambiguation on which some rights are in fact basic; and §3 examines Andrew Cohen's argument that subsistence in particular is not a basic right, and argues that he, too, fails to give adequate attention to the normative element of a right.

1 Shue's argument for basic rights

Shue defines a moral right as:

- (1) the rational basis for a justified demand that
- (2) the actual enjoyment of a substance be
- (3a) socially guaranteed against
- (3b) standard threats.¹

¹Shue [5] p. 13. I split condition (3) into (a) and (b) for ease of exposition.

The normative aspect of (1) is crucial. What distinguishes a right from a claim of other kinds is that the demands in (2) and (3) are justified: ‘good reasons’ support the claim that the guarantees for some substance ‘ought to be provided’.² Shue takes these guarantees (3a) to consist in certain duties, so a necessary feature of a right, according to this definition, is that there are normative considerations weighty enough to justify imposing these duties for the sake of guaranteeing some substance (2) against standard threats (3b).³

The ‘substance’ in (2) is the object to which the right is a right. That this is to be ‘actually enjoyed’ is intended to establish that the sense in which it must be possible to exercise a right is not merely logical, but must be such that the right-holder’s interest in exercising the right is not outweighed by the threat of harm, since Shue rejects any notion of enjoyment ‘compatible with being unable to make any use of the substance of the right.’⁴

Similarly, simply enjoying the object of a right is not by itself sufficient for one to have a right to that object. The idea is that it continues to be possible that the right-holder could be effectively prevented from enjoying the object of their right in the absence of social guarantees (3a), cashed out in the form of duties imposed on certain people and institutions, even if they are at present enjoying the right’s object.⁵ It is not until arrangements that effectively prevent people from depriving others of the object of their right are in place that one actually enjoys a right to that object.⁶ These guarantees are not absolute; they ensure only that one’s ability to enjoy the right is reasonably secure.⁷ What counts as reasonable will be a question determined both by the normative considerations supporting the right in the first place, and by empirical circumstances.⁸ If the right is justified because it protects a fundamental interest, the same considerations cannot simultaneously sanction social guarantees so costly that they negatively affect the very same interest of whoever is responsible for discharging the duties correlated to the right. The costs determining reasonableness will therefore be sensitive to the empirically

²Ibid. p. 13.

³Ibid. p. 16.

⁴Ibid. p. 20.

⁵Ibid. p. 22.

⁶Ibid. pp. 16–17.

⁷Ibid. p. 17.

⁸Ibid. p. 33.

determined abilities of the duty-bearer to discharge that duty.⁹

The threats against which the guarantees in (3a) protect are limited by (3b) to ‘standard threats’, rather than all circumstances incompatible with the object of a given right.¹⁰ Standard threats to the enjoyment of (the object of) a right are ordinary and serious, but within society’s ability to remedy.¹¹ Shue’s paradigm examples are those conditions that are either directly incompatible with, or which could be used coercively to prevent people from enjoying, the substance of some right; but this latter category includes any “serious [...] hinderance to the enjoyment of any rights”.¹² Thus, standard threats to a right against assault are both being assaulted or being coerced into being assaulted, though the latter is but one example of the range of things that could be expected to prevent one from enjoying the right’s object.

Thus, to claim a right to *X* is to claim that there are (1) normative considerations weighty enough to justify (2) the enjoyment of some substance being (3a) guaranteed to a reasonable degree against (3b) ordinary and preventable threats that either directly conflict with enjoying the object or which could be expected to hinder the enjoyment of that substance. Note that the guarantees a right provides are therefore limited in two respects: they extend only to a subset of the things that could threaten one’s enjoyment of (2) and provide only a reasonable degree of security against such threats. These aspects differ in their sensitivity to empirical circumstances. Roughly: standard threats are sensitive to what is around to threaten one’s enjoyment of the object of a right; the reasonableness of guarantees is sensitive to the duty-bearer’s capacity to do something about these threats at a reasonable cost.

Shue sees basic rights as following analytically from the concept of a moral right, that is, simply in virtue of the meaning of the definition.¹³ They are also the absolutely minimal moral claims of every person on everyone else.¹⁴ His formal argument proceeds as follows:

- i. Assume there is at least one thing, *X*, to which everyone has a right, where a right is defined as above;

⁹Ibid.

¹⁰Ibid. p. 29

¹¹Ibid. p. 33, 25.

¹²Ibid. p. 21. Both of Shue’s examples are framed in these terms, see pp. 20–21, pp. 22–24.

¹³Ibid. p. 31.

¹⁴Ibid. p. 19.

- ii. Some other things are necessary for enjoying *X as a right*; therefore,
- iii. ‘Everyone also has rights to those other things that are necessary for enjoying [*X*] as a right’.¹⁵

Since enjoying *X* as a right is to enjoy it to the degree provided for by the guarantees in (3), the italicised qualification in (ii) limits the range of things that are necessary to those in the scope of those guarantees, rather than letting (ii) encompass anything necessary for enjoying *X* itself. This curtails the list of rights that count as basic to those that “constitute a standard threat to all other rights”.¹⁶

How does this formal argument relate to the normative element of rights? Notice that the normative considerations in (1) will be partly determinative of what the object of the right is. Shue’s definition is neutral as to what considerations actually justify a given right, but to illustrate, let (1) be a simple Utilitarian calculus according to which goods are ranked by the amount of (dis-)pleasure they produce. Certain goods under this framework will tend to regularly outweigh others, e.g. the amount of displeasure one would incur from being assaulted. That a right justifies a demand points up the fact that only weighty interests are the kinds of thing apt to be the object of a right.¹⁷ Regardless of the substantive specification of (1), any substance that would constitute a standard threat to enjoying that non-basic right is likely to be among our most fundamental interests.¹⁸ Any normative considerations that justify demands for some non-basic right will therefore justify a right to one of these basic goods at least as strongly, because our interest in these goods will give them huge weight in something like a utilitarian calculus. There is thus a transitive relationship between the justification for the non-basic right and the things necessary for enjoying that right: if a right-holder *P* has a right to *X*, and *Y* is necessary for enjoying a right to *X*, *P* has a right to *Y*.¹⁹ No consistent set of normative considerations could justify a right to *X* without also justifying a right to *Y*. Shue’s claim is that subsistence is just such a fundamental

¹⁵Ibid. p. 31.

¹⁶Ibid. p. 19.

¹⁷The possibility of other considerations outweighing those adduced to support (2) and (3) is left open by the definition. Ashford [1] p. 95 notes that this relationship of weights between interests is what grounds the interdependence between rights.

¹⁸Ashford [1] makes a similar point, although she is not appealing specifically to a utilitarian calculus.

¹⁹Ibid. pp. 32–33.

interest, and therefore lacking (a right to) subsistence is a standard threat to all other rights.

Shue's claim that basic rights follow by analytic necessity means he must defend two claims:

General It is impossible to enjoy a right to X without also enjoying basic rights, whatever these basic rights are;

Particular It is impossible to enjoy a right to X without also enjoying (a right to) subsistence.

The remainder of this essay considers challenges to these claims in turn.

2 Pogge's critique of Shue

Thomas Pogge offers three principal construals of what is meant by 'enjoying a right' in premise (ii) in order to understand how basic rights are necessitated by this enjoyment. He claims that each of these is open to counterexample, and his argument therefore takes aim at the general claim above.²⁰ Pogge defines a notation that relates each part of Shue's definition of a moral right to a person:²¹

$A(PX)$ = ' P actually enjoys X ';

$B(PX)$ = ' P 's enjoyment of X is socially guaranteed against standard threats';

$C(PX)$ = ' P enjoys X as a right'.

$A(PX)$ and $B(PX)$ correspond to (2) and (3) respectively; $C(PX)$ is alleged to correspond to (1). Pogge takes these conditions as individually necessary and jointly sufficient for a person P to have a right to X .²² He then considers the following entailments that might hold between P 's right to some non-basic substance Y and a candidate basic X , and concludes each is open to counterexample:

(a) $A(PY) \rightarrow A(PX)$ (b) $AC(PY) \rightarrow AC(PX)$ (c) $ABC(PY) \rightarrow A(PX)$

Two points require immediate attention. First, Pogge's rendering of part (1) of a moral right is peculiar. He interprets Shue as advocating three conditions

²⁰Pogge [4] p.117.

²¹Ibid. pp.114–115.

²²Ibid. p.114.

constitutive of $C(PX)$: “the contributions others must make to $[A(PX)]$ and $[B(PX)]$ being met are understood as duties (C_1) that are owed to P (C_2); and $[C_1$ and $C_2]$ are culturally anchored (C_3)”.²³ Failing to satisfy any of C_1 – C_3 is sufficient for $C(PX)$ to fail to be satisfied. The problem with this is that $C(PX)$ only partially captures the normative force of (1). Setting C_3 aside, it seems clear that C_1 and C_2 could be entailments of whatever normative considerations appear in (1), but (1) is also partly determinative of the substance of the right, a feature that C_1 and C_2 fail to capture.²⁴ Moreover, the social guarantees in $B(PX)$ entail Shue’s tripartite schema of duties to avoid, protect and aid.²⁵ If $C(PX)$ is meant to cover ‘what $[A(PX)]$ and $[B(PX)]$ leave out’, it cannot include C_1 and C_2 , since these are already part of $B(PX)$.²⁶

Second, $A(PX)$ is not in fact a necessary condition for enjoying a right, *contra* Pogge. One might be deprived of the substance of a right as a result of some threat to which social guarantees do not (reasonably) extend. Since Shue is arguing that if one enjoys Y as a right one must enjoy a right to X , and a right to Y does not entail $A(PX)$, interpretation (a) can be dismissed as irrelevant; (b) can be dismissed on similar grounds, since it omits the crucial social guarantees that are the centrepiece of Shue’s definition of a right. Before (c) can be evaluated it requires revision. We can revise $C(PX)$ to capture the normative force of (1):

$C^*(PX) =$ ‘There are good normative reasons for P to be able to enjoy X ’

$A(PY)$ can also be dropped from the formulation of what enjoying a right consists in. It is only necessary for enjoying a right that P be able to enjoy its substance with reasonable security, and this latter condition is captured by $B(PY)$. Thus enjoying a right is captured by:

$$(c^*) BC^*(PY) \rightarrow A(PX)$$

But is this correct? The problem is that the sense in which enjoying a basic right is necessary is ambiguous, notwithstanding Shue’s claim that he means ‘made necessary by the very concept of a right’.²⁷ Shue’s claim simply narrows the range of things that are entailed by non-basic rights to reasonably

²³Ibid. p. 15.

²⁴ C_3 is contentious as a necessary condition on normativity, and is not required for Shue’s argument.

²⁵Shue [5] p. 52.

²⁶Pogge [4] p. 115.

²⁷Shue [5] p. 31.

preventable standard threats (3a, 3b). But this is compatible with two interpretations of (ii) between which Shue vacillates: enjoyment of a right consisting in (α) having the *object* of that right; and (β) having the social guarantees of the right. This corresponds to two different analyses of (ii), according to which the reason basic rights are entailed by non-basic rights is: (α) their objects are necessary for enjoying the non-basic right, i.e. (c^*); and (β) whatever social guarantees are necessary to safeguard the object of the basic right are necessary to safeguard the non-basic right, i.e. (c'):

$$(c') BC^*(PY) \rightarrow B(PX)$$

Both of $\neg A(PX)$ and $\neg B(PX)$ are standard threats, since either could hinder one's enjoyment of a non-basic right, and both (c^*) and (c') are therefore compatible with (ii) being a claim of conceptual necessity. What is at issue is which of these Shue requires in order to substantiate the claim that enjoying a right to basic X is necessary for enjoying a right to non-basic Y .²⁸ This distinction is important, for it directly affects the kind of counterexample one must give in order to disprove Shue's thesis.

Shue claims to take 'enjoying a right' to be an elliptical way of saying P enjoys the substance of that right.²⁹ But recalling that Shue does not think it is necessary for having a right that one enjoys its substance at all times, this cannot be right. To show P lacks a basic right, it does not suffice to show that there is some possible state of affairs in which $BC^*(PY) \wedge \neg A(PX)$, since it might be the case that the deprivation of the object of the basic good in question is in fact the result of a non-standard threat or one of the small number of deprivations allowed under the prevailing standard of reasonableness. Both of these features are captured by the consequent of (c'), so (c') must be the correct interpretation of what enjoying a right consists in. What Pogge needs to show to defeat the general claim, then, is that there is a situation in which there are good normative reasons for P to have X , and that P has all the social guarantees for the non-basic good Y , yet at the same time fails to have guarantees for a basic right(s). Can Pogge show there is a situation of this kind? It is not clear that he can. Pogge offers the example of the right not to be arbitrarily deprived of one's nationality (Y), and claims that all of $ABC(PY)$ obtain with respect to Y , without it also being the case that one enjoys subsistence (X).³⁰ In order

²⁸Other possible interpretations are $BC^*(PY) \rightarrow A(PX) \vee B(PX)$, which would be too inclusive; and $BC^*(PY) \rightarrow A(PX) \wedge B(PX)$, which is too exclusive.

²⁹Shue [5] p. 15.

³⁰Pogge [4] p. 121.

for Shue to claim P lacks a right to Y , says Pogge, he would have to “assert that it is never possible to fully enjoy [a right to Y] without enjoying [...] subsistence”.³¹ Pogge then posits a hypothetical situation in which P is kidnapped and is deprived of subsistence. In this situation, P would be ‘willing to give up his nationality to end his ordeal. But this outcome is evidently of no interest to his captors [...] [so there] is no reason to doubt that P continues to enjoy fully his right to [Y].’³² The idea is that the scenario $ABC(PY) \wedge \neg A(PX)$ obtains because no one is actually attempting to coerce P into giving up his nationality in exchange for subsistence, so P continues to enjoy Y , without enjoying X . This makes $ABC(PY) \rightarrow A(PX)$ contingent on whether anyone is actually interested in coercing P into foregoing their non-basic right in exchange for a basic one, rather than a necessary entailment of enjoying a right to Y . So (ii) is false. Ashford offers a reply on Shue’s behalf that trades on the weight of various interests. Subsistence (and other basic rights) has two features: lacking it is ‘unsustainable’, and it is therefore likely to outweigh P ’s interest in the substance of any other right.³³

Because of these features, the threat to enjoying a right to non-basic Y is not that someone actually is interested in coercing P into foregoing Y , but that in any scenario in which P lacks access to subsistence, if this were to conflict with the exercise of another right, P would not exercise the non-basic right, because P ’s interest in doing so would be outweighed by their interest in access to subsistence. Person P ’s refraining from exercising the non-basic right in this scenario would be just as involuntary as when someone is actually coercing P .³⁴ So the possibility of lacking secure access to subsistence, which exists in the absence of social guarantees, is itself a standard threat. This does not mean that at all times P must actually enjoy the substance of a basic right – P might voluntarily give it up – but simply that, if P is to be able to exercise a non-basic right, then guarantees must exist for the basic right such that P ’s decision to exercise their non-basic rights is not constrained by the possibility of lacking access to subsistence, since the latter is always of greater importance. This explains the entailment that is expressed by the analysis of (ii) as (c’). Therefore, Pogge’s objection fails; P does require guarantees to subsistence to enjoy a right to nationality. Moreover, the reply offered here specifies a relation

³¹Ibid. pp. 120–121.

³²Ibid. p. 121

³³Ashford [2] p. 501.

³⁴Ibid. pp.501–502.

that holds between the substances of all rights, namely, an ‘objective’ weighing relation.³⁵ Because of its generality, it can do the work of buttressing the claim in (c’): no right can be enjoyed without also enjoying social guarantees to non-basic rights, and this is because of the relationship that obtains between *P*’s varying interest in the substances of different rights.

3 Cohen’s torture contract

Andrew Cohen appears to accept the analysis of the necessity claim as (cc’), but objects that the particular duties that attach to subsistence rights are not required to enjoy other rights, thus taking aim at the particular claim from §1. Shue’s definition of a moral right, says Cohen, “necessarily imposes some positive duties”.³⁶ By giving an example in which a non-basic right is enjoyed without also enjoying the positive duties Shue claims are imposed by subsistence, Cohen aims to set up a dichotomy for Shue: either basic rights do not all impose positive duties, a central feature of Shue’s attempt to show parity between so-called negative rights and subsistence rights; or subsistence is not a basic right (Cohen 2004, p. 266-7).³⁷ His example is an evolution of an objection originally considered by Shue in a footnote, in which *P* enjoys freedom from torture, including social guarantees against torture (as Cohen notes, fulfilled by discharging negative duties not to torture, positive duties to protect *P* from torture and aid *P* when tortured), yet the positive duties to *P* that are meant to be partly constitutive of the social guarantees attending a right to subsistence have not been discharged.³⁸ Cohen aims to counter Shue’s original riposte, that *P* is always suffering a standard threat to his right against torture because vulnerable to exchanging this right for subsistence, by building effective guarantees against this kind of contract into his example so that *P* is never in fact coerced, and therefore still enjoys the right against torture without enjoying the right to subsistence.³⁹ This would be so even if we employ Ashford’s response to Pogge, because where her reply essentially allows that the possibility of trading a non-basic right for subsistence is a standard threat because one would virtually always accept this trade, Cohen’s objection rules

³⁵Ibid. p.501.

³⁶Cohen [3] p. 265

³⁷Ibid. pp. 266–267

³⁸Ibid. p. 268

³⁹Ibid.

this possibility out by hypothesis.

Cohen's objection looks powerful. However, like Pogge, Cohen overlooks the vital part (1) of Shue's definition of a moral right in the following way. Recall the analysis of the necessity claim in (ii), (c'). According to this analysis, enjoying a right consists in there being good normative reasons for *P* to enjoy *Y* and social guarantees for *Y*. Recall also from §1 that these normative reasons justify those guarantees. If the entailment in (c') holds, the normative considerations that justify a right to non-basic substance *Y* also justify the guarantees to the basic substance *X*.⁴⁰ So Cohen would have to show that there is some set of normative considerations that justifies social guarantees against torture and torture contracts, guarantees which include positive duties, yet at the same time fail to justify social guarantees for subsistence. But what possible set of normative considerations could do this? It is not obvious there are any, and the reason is related to Ashford's point about objective interests: any plausible normative principles cannot be divorced from the interest of rights-holders.⁴¹ In fact, as discussed in §1, they partly determine what is apt to be the object of a right on the basis of rights-holders' interests. Given this, and the weightiness of subsistence, they could not possibly justify $B(PY)$ without also justifying $B(PX)$. This is not a substantive moral claim, but rather a minimal constraint on the meaning of 'justified' in (1): the normative considerations at issue must at least be sensitive to the interests of the rights-holder, so Shue's claim remains conceptual. At best, then, Cohen has shown that the social guarantees to *Y* do not entail all the social guarantees to *X*, i.e. that $\neg B(PY) \rightarrow B(PX)$. But this conditional is not what is at issue in Shue's necessity claim; (c') is. Cohen's claim that we can enjoy non-basic *Y* without also enjoying the positive duties entailed by social guarantees to basic *X* is false, because in such a case we would no longer be enjoying *Y* as a right, but rather a brute set of social guarantees divorced from plausible normative considerations that justify those guarantees as rights. I conclude that Cohen's dichotomy can be resisted. Person *P* cannot enjoy a non-basic right without this entailing the positive duties that guarantee subsistence. His objection to the particular claim in §1 therefore fails.

⁴⁰See §1. The justificatory relationship is transitive between non-basic and basic rights.

⁴¹Ashford [1] p. 99.

4 Conclusion

On a construal of Shue's claim that non-basic entail basic rights as (c'), both the general and particular claims in §1 can be defended. The basis for their defence is an entirely general consequence of the analysis of the necessity claim (ii) as (c'), since the normative considerations justifying any non-basic right could not fail to justify protecting a fundamental interest like subsistence. This rules out the kinds of counterexample that could serve to undermine either the general or particular claim. Shue succeeds in showing subsistence is a basic right in that his argument is at least valid. Critics must therefore attack something other than the non-basic – basic right entailment in order to reject his conclusions.

References

- [1] Ashford, E. (2009) 'The Alleged Dichotomy between Positive and Negative Rights and Duties', in C.R. Beitz and R.E. Goodin (eds.), *Global Basic Rights*. Oxford: Oxford University Press.
- [2] Ashford, E. (2009) 'In what Sense is the Right to Subsistence a Basic Right?', in *Journal of Social Philosophy* 40: 488–503.
- [3] Cohen, A. (2004) 'Must Rights Impose Enforceable Positive Duties?', in *Journal of Social Philosophy* 40: 264–276.
- [4] Pogge, T. (2009) 'Shue on Rights and Duties', in C.R. Beitz and R.E. Goodin (eds.), *Global Basic Rights*. Oxford: Oxford University Press.
- [5] Shue, H. (1996) *Basic Rights: Subsistence, Affluence, and US Foreign Policy* (2nd edn.). Princeton, N.J.: Princeton University Press.

A Set of Objections to David Benatar's Anti-Natalism

Franco Palazzi

University of Pisa

1 Introduction

David Benatar has recently argued that coming into existence is always a harm and that therefore there are compelling moral reasons in favor of a gradual extinction of humankind.¹ Several authors have replied to his core thesis, but they frequently dealt with only some parts of his work on anti-natalism, often rejecting it altogether in a somewhat precipitous way which allows for – at times convincing – replies.² In my view, even the profoundest critics failed to recognize some fundamental mistakes Benatar made in his provocative defense of anti-natalism. In this article, I try to take into account *all* of Benatar's main arguments, raising several objections to each of them and also developing, in paragraph four, an original view of the meaning of life and its relationship with worthwhileness.

2 Benatar's main arguments

Usually both scholars and laypeople assume the intrinsic negativity of human extinction in a scenario in which, as in the present one, life can be judged worth-living. Nevertheless, the fact that human life could be – and often is, today – worth-living is disputable. Recently, David Benatar has strongly argued that it would be better if humanity became extinct because coming into existence is *always* a serious harm.³ Although almost none (at least amongst professional philosophers) seem to endorse his provocative view, its thought-

¹Benatar [2].

²Benatar [3].

³Benatar [2]. This is not to suggest that the very word “anti-natalism” means that we should implement extinction as soon as we can. Indeed, for Benatar, as well as in general, this term indicates the fact that procreation is wrong.

fulness and frequent presence in academic debates drive me to analyse – and criticise – it in some detail.

According to one of Benatar’s main claims, called ”Basic Asymmetry”, pain is bad and pleasure is good when they are present, but when they are not, the absence of pain is good (even if it is not enjoyed by anyone) whereas the absence of pleasure is not bad (unless there is somebody who is deprived of it).⁴ Put in another way: “judged in terms of the interest of a person who now exists, the absence of pain would have been good even though this person would then not have existed”.⁵ So, consistent with the interests of ideally possible people, absence of pain is good but absence of pleasure is not bad as it can’t be experienced by anyone who doesn’t exist yet, therefore it would be better if no human being existed

Such an argument, Benatar goes on, explains four other ‘asymmetries’: while there is a duty to avoid bringing suffering people into existence, there is no duty to bring happy people into being; whereas it is strange to give as a reason for having a child that the child will thereby be benefited, it is not strange to cite a potential child’s interest as a basis for avoiding bringing her into existence; only bringing (rather than not bringing) people into existence can be regretted for the sake of the person whose existence was contingent on our decision; while we are rightly sad for suffering foreign people, we are not similarly sad for the happy people who, had they existed, would have populated a desert island.⁶ This is Benatar’s First Argument.

One obvious objection to this view says that, if the great majority of us still thinks that life is worth-living, Benatar’s position is absurd. Predictably, he anticipates this criticism and tries to demonstrate its inconsistency using a double *escamotage*. Firstly, he repeats several times that affirming that human life is never worth-starting doesn’t necessary imply, once life has started, the desirability of death – and so of suicide. “This is because the existent can have interests in continuing to exist, and thus harms that make life not worth continuing must be sufficiently severe to defeat those interests”, whereas “the non-existent has no interest in coming into existence”.⁷ Indeed, “[c]oming into existence

⁴Note that Benatar is not a hedonist, because he uses pleasure and pain only as *exemplars* of two broader categories: benefit and harm. Benatar [2] p. 30; [3] p. 122 n. 6.

⁵Benatar [2] p. 31.

⁶Ibid. pp. 31–35.

⁷Benatar [2] p. 213.

is bad in part because it invariably leads to the harm of ceasing to exist”.⁸ Secondly, he claims that self-assessments of one’s quality of life are psychologically unreliable due to phenomena like Pollyannism (a tendency towards optimism which, for example, inclines us to recall positive rather than negative experiences or to believe that we are better off than the average person), adaptation (the capacity to adapt to a new situation and to adjust one’s expectations accordingly) and implicit comparison (because of which self-assessments are better indicators of relative rather than absolute well-being) (Benatar 2006, pp. 64–69).⁹ This is the core of Benatar’s Second Argument.¹⁰

In the following paragraphs I will try to show why we should reject at all his anti-natalism.

3 The Basic Asymmetry: some preliminary difficulties

Benatar’s Basic Asymmetry has a fragile theoretical basis. Indeed, if the absence of pain is good even if it is not experienced by anyone, there are no logical reasons to deny that, symmetrically, the absence of pleasure is bad – even if nobody experiments it.¹¹ At least, Benatar should be required to convincingly demonstrate the existence of a wider – and improbable – asymmetry between bad and good in order to defend his original claim.¹² Unfortunately, he has not provided such a demonstration. However, he has replied to this by saying that his argument is not logic, but axiological; that is, we *can* affirm, from a logical standpoint, that the absence of pleasure is bad *per se*, but we *should* not. According to him, the most prominent motive to maintain Asymmetry is that it provides, in a simple and homogeneous way, an explanation of the other four asymmetries¹³ as well as a solution to a set of Parfitan problems

⁸Ibid.

⁹Ibid. pp. 64–69.

¹⁰It seems that Benatar considers his two Arguments both independent and interrelated at the same time: on the one hand, while the First is aimed at showing that life is always a harm for the one who lives, the Second tries to show that this harm is usually a notable one; on the other hand, the Second Argument can also be understood as a separate argument for the conclusion that coming into existence is a harm? (Benatar [3] pp. 123–124; see also Benatar [2] p. 14). As I will reject both of them, I think that their precise relationship is indifferent for the purposes of this article.

¹¹Harman [10] pp. 781–782.

¹²Ibid. pp. 785 n. 3; McMahan [16] pp. 63–64.

¹³Benatar [3] pp. 126–127.

in population theory like the Repugnant Conclusion.¹⁴

It seems to me that, in giving this ingenious reply, Professor Benatar has committed several argumentative mistakes.

The first consists in implicitly believing that, once an asymmetry between *pleasure* and *pain* is accepted, it could be *ipso facto* extended to the broader level of *benefit* and *harm* – of which pleasure and pain are only exemplars (see note 3 above). However, I will say more about this in the paragraph on the ‘structure of meaning’.

The second, more evident mistake is an epistemological one: the ability of a theory to explain more phenomena than its competitors could be a good argument in physics, but it is much more controversial in ethics for a number of reasons. First of all, it is not taken for granted that there are, in this context, phenomena to elucidate at all: Benatar’s four other asymmetries are not based on philosophical arguments, but on moral intuitions assumed as widely spread. I have the impression that these intuitions are not so widespread (at least among non-philosophers) – and some recent studies in folk psychology show that often philosophers are somewhat precipitous in presuming the average person’s intuitions.¹⁵ Even acknowledging that the asymmetries at stake are largely intuitive, their philosophical correctness is not obvious (in particular, I will argue that the fourth is not an asymmetry at all in the last paragraph of this article). A similar point can be made about Parfit’s puzzles in population ethics. For example, some authors have claimed that the Repugnant Conclusion¹⁶ is not a consequence of total utilitarianism¹⁷ or that it could not be repugnant whatsoever.¹⁸

But suppose that Benatar is able to demonstrate that his Basic Asymmetry answers real and significant questions. He would still have to show why these answers are better than others’.

This reflection leads us to the third mistake. Benatar often seems to lessen alternative positions with the argument that his is preferable because it can solve

¹⁴Benatar [2] pp. 172–182.

¹⁵See, for example, Nahmias, Morris & Nadelhoffer [20] on intuitions about moral responsibility and determinism.

¹⁶Parfit [23] pp. 387–390.

¹⁷e.g. Hare [9] pp. 72–83.

¹⁸Huemer [12].

a considerable amount of heterogeneous problems which otherwise would require several *ad hoc* solutions.¹⁹ This is a very risky assumption: let us consider a couple of counterexamples. Imagine that two detectives are investigating a series of murders in the same town. Through parallel investigations they formulate different hypotheses: according to the first there are two separate culprits, whereas the second thinks that the only responsible is a third man. Now, suppose not to know further details about the two hypotheses or the murders: it would be very improbable to say that, for the sake of simplicity and completeness, the second detective is right while her colleague is wrong. Benatar could object that the criteria used in evaluating reality are different than those required to defend a moral theory; so maybe we need a counterexample taken from a more philosophical context. If so, take Hegel's system: it gave us a single and coherent view of a great variety of philosophical themes, from the philosophy of nature to religion, using a schematic triadic structure and solving a lot of theoretical problems. Nonetheless, nobody seems to endorse it – except, maybe, for some of its parts. I do not want to deny that in philosophy simplicity is often a quality, but simplicity per se is not a good reason to reject more intricate arguments, nor a preliminary advantage Benatar's thesis has over different, somewhat more complicated ones.²⁰

4 The Second Argument: an unacceptable kind of pessimism

Arguing in favour of the unreliability of life-assessments of one's quality of life, Benatar stresses the relevance of positive memory bias and adaptation (see above). However, I will demonstrate that his use of the psychological literature on this phenomena is involuntarily partial, and leads to conclusions that, at least in his formulation, are not acceptable as philosophically relevant.

Starting with the memory bias, Benatar seems not to acknowledge a wider tendency according to which events that are negatively valenced have a greater impact on individuals than positively valenced events of the same type.²¹ For example²² we adapt more rapidly to a positive rather than to a negative change

¹⁹e.g. Benatar [3] p. 127.

²⁰However, in paragraph 6, I will try to formulate a simple and homogeneous alternative solution of Benatar's first three asymmetries.

²¹Baumeister *et al.* [1].

²²I owe the following examples (and references) to *ibid.*

in our lives;²³ a negative day influences accordingly one's wellbeing the day after, while a positive one does not;²⁴ within marital life, negative events have more consequences than their positive equivalent;²⁵ pessimism and negative events condition one's health more than optimism and positive events.²⁶ In this context, the positive memory bias is only an exception – so we cannot argue that our minds are blindly imbalanced toward optimism.²⁷

Furthermore, memory is a more complex mechanism than what Benatar tells us. It has been observed that negative events evoke strong and rapid psychological, cognitive, emotional and social responses. Such a mobilisation is followed by psychological, cognitive and behavioural responses that minimise and sometimes even erase the impact of those events. This mobilisation-minimisation pattern is greater for negative events than for neutral or positive ones.²⁸ Recently, Ritchie *et al.* have found that after one year emotions associated with recalling negative autobiographical events fade in 51% of the cases, whereas emotions associated with remembering positive events fade in 37% of the cases: a Fading Affect Bias (FAB) certainly exists, but fading affect is not a one-way phenomenon.²⁹ Moreover, FAB is particularly significant in the short term, while in the long time it is weakened by other cognitive phenomena.³⁰ In a similar way, Taylor and Armor recognised that positive illusions (namely self-aggrandisement, unrealistic optimism and exaggerated perceptions of control) typically remain within modest bounds.³¹ This is probably because memory biases and other positive illusions are adaptive,³² helping us not only to maintain good subjective well-being levels, but also to be better off in respect to some objective criteria (physical and mental health, concentration, social relationships). If they were too strong, they would become maladaptive, making us lose the positive effects. On a wider level, there are good reasons to think

²³See Brickman *et al.* [4].

²⁴Sheldon *et al.* [26].

²⁵See Hunston & Vangelisti [13].

²⁶Baumeister *et al.* [1] p. 353.

²⁷*Ibid.* p. 344.

²⁸Taylor [27].

²⁹Ritchie *et al.* [25].

³⁰Walker *et al.* [29].

³¹Taylor & Armor [28].

³²See Taylor [27] and Walker & Skowronski [30].

that also the psychological superiority of bad over good is a kind of adaptation.³³ Finally, although this prevalence of bad, the majority of people declare good levels of wellbeing. We cannot exclude that such result is caused not only by positive illusions, but also by the fact that many lives contain far more good than bad.³⁴

Now, I want to consider adaptation. This is traditionally a very controversial issue due to its political implications: members of disadvantaged categories often adapt their preferences in accordance with those offered them by the structure of society, which usually incorporates some kind of imbalances of power and, using Bourdieu's term, "symbolic violence".³⁵ Nonetheless, this does not imply that all kinds of adaptive preferences³⁶ are *per se* negative. Once assumed, following Nussbaum, that we should not consider as authentic preferences not to have, for example, political liberties and rights, or literacy, we have to notice that adaptive preferences often drive us to make positive choices.³⁷ Indeed, they help us not to focus too much on our failures and misfortunes, prompting us to develop an autonomous and not defeatist vision of our selves and at the same time often increasing our subjective wellbeing.³⁸ In short, there could be good reasons to adopt adaptive preferences – even though, of course, these reasons could be outweighed by others.³⁹ According to David Bruckner, we can distinguish between rational and irrational adaptive preferences by recognising the former as objects of the agent's *reflective endorsement*. This term means not a preference for the preference already chosen, but an overall judgment able to conflict with second-order preferences.⁴⁰

In addition to the fact that adaptation is not always bad, or irrational, it is important to consider the context which Benatar criticises this widespread tendency in: he makes reference to cases in which "a person's objective wellbeing takes a turn for the worse" and the consequent dissatisfaction is soon followed

³³Baumeister *et al.* [1] pp. 357–358.

³⁴*Ibid.* pp. 361–362.

³⁵See, for example, Nussbaum [22] and Levy [15]. about gender-adaptive preferences.

³⁶Here I follow Bruckner [5], in considering that an adaptive preference change "occurs when an agent's preference changes as a result of a change in the agent's feasible set [of options]". (p. 308).

³⁷Nussbaum [22] pp. 83–84.

³⁸Bruckner [5] pp. 312–314.

³⁹*Ibid.* p. 315.

⁴⁰*Ibid.* p. 317–319.

by a propensity “to adapt to the new situation and to adjust one’s expectations accordingly”.⁴¹ Clearly enough, he describes an autonomous kind of adaptation that does not convey the negative influence of an asymmetry of power. It seems to me that this is a positive kind of adaptation, maybe one which enables a person to live well despite her hardships.

It could be objected that, even if adaptation really has the power to increase a person’s wellbeing to some extent by encouraging that person to feel good despite her notable sufferings, this in any case would not be enough – in other words, she will never be so good as she deceives herself to be, not precluding herself from radical evaluative errors.⁴² But this is a weak reply. First, Benatar refers to a pattern of physiological dynamics, leaving aside pathological cases; however, as we have already seen, the psychological illusions he mentions remain normally within modest limits. Second, these dynamics are not the only ones we experiment: if human beings systematically ignored suffering or sadness, nobody would be depressed. Third, by raising this criticism, Benatar seems to make reference to a sort of universal and almost scientific notion of wellbeing according to which an individual, although autonomously (and not pathologically) convinced of her wellbeing, could anyway be in a very bad – absolute – state. This move is dubious both empirically and theoretically. From an empirical standpoint, we should take note that many economists believe that subjective wellbeing is a good parameter when measuring happiness. Interestingly, two scholars, skeptic about its appropriateness in economics, developed a new index expressly aimed at avoiding biased dynamics such as adaptation.⁴³ By adopting their index, however, we find that most people have quite high levels of wellbeing. Turning to theoretical aspects, denying the authenticity of people’s self-assessments can easily lead Benatar to paternalistic conclusions. For example, if an autonomous person usually is not able to give an evaluation of her quality of life which is worthy of consideration, how can she give reliable evaluations at all? Can she be considered free in the philosophical meaning of the term? I think that the answer should be probably a negative one: if a person is not able, except by making a huge philosophical and psychological effort,⁴⁴ to acknowledge the *quantity* of bad and good fea-

⁴¹Benatar [2] p. 67

⁴²Benatar [3] p. 146.

⁴³Kahneman & Krueger [14].

⁴⁴Some people may think that, through his writings, Benatar wants indeed to render this effort a very easy one, suitable for almost all of us. But he seems disenchanting in this respect: “[i]t is unlikely that many people will take to hearth the conclusion that coming into existence is

tures present in her life, how can she clearly distinguish their *quality*? That is, how can she definitely differentiate pleasure from pain, benefit from harm, and so on? A bit more radically: how can she have free will? If Benatar wants to maintain his pessimistic view of human psychology, he will have to explain how his theory could avoid paternalistic conclusions.

A final thing must be said before concluding the current paragraph. Whereas Benatar argues that the Second Argument can also be understood as a separate argument for the conclusion that coming into existence is a harm (see note 4), the same psychologists he mentions seem not to endorse this view. To my knowledge, none of them assumes that, say, the Pollyanna Principle implies anti-natalism. Perhaps Benatar thinks that his interpretation of their findings is better than their own, but if so, he should reasonably provide a damning evidence that, at the moment, he has never shown. For all these motives, I conclude that his Second Argument is based on an unacceptable kind of pessimism.

5 The (Asymmetric) Structure of Meaning

In the second paragraph of this article I notice that Benatar implicitly tends to extend his Asymmetry between *pleasure* and *pain* to the broader level of *benefits* and *harms*. Now I will claim that this argumentative shift is mistaken because it induces us to overshadow the great importance that *meaning* has in every life. Furthermore, I will try to show that the meaning of life is asymmetric in a way which questions the truth of Basic Asymmetry. As the reasoning I have in mind is a bit complex, I will firstly provide a few definitions of and distinctions between some concepts used throughout this paragraph.

Primarily, we have to distinguish amongst the notions of “happiness”, “worthwhile (or worth-living) life” and “meaningful (or significant) life”. With the first term I mean a mental state identified with an agent’s emotional condition as a whole.⁴⁵ The second refers to a life whose subject chooses to continue after an autonomous reflection (i.e. a reflection carried out in a context not

always a harm. It is even less likely that many people will stop having children. By contrast, it is quite likely that my views either will be ignored or will be dismissed. [...] this response [...] does result from a self-deceptive indifference to the harm of coming into existence” (Benatar [2] p. 225). I think that such a statement gives even more power to the objection according to which Benatar’s view is (aristocratically) paternalistic.

⁴⁵Haybron [11] §2–2.4.

affecting her preferences in an oppressive way, nor distorting them in such a way that they would no longer be considered hers). The third term is related to an existence that, “largely in virtue of one’s actions and their causes and consequences, warrants great pride or admiration, or exhibits superlative final goods beyond one’s animal self”.⁴⁶ More precisely, a life can be meaningful in two ways: autobiographical and biographical. The former implies that I consider my life meaningful, usually as part of a relationship with a wider context exceeding the limits of my person (it could be a moral or political ideal, a net of affective relationships – or even a single relationship – a religious faith or simply devotion to art). The latter, on the other hand, deals with a third-person judgment on the meaning of someone else’s existence. Though autobiographical and biographical meaning could often coincide (at least in a general manner), it is useful to stress some differences: while the first must obviously be recognised within the limits of one’s own life, the second can be attributed to an already ended existence (which better applies to an overall evaluation); the autobiographical meaning focuses more on the reasons motivating one’s actions, whereas the biographical meaning pays greater attention to their effects. One level frequently influences the other: it’s easier to consider one’s life meaningful if many people believe that it is so, and a person who is strongly convinced of living a meaningful existence has more of a chance to persuade other people of this fact than one who is perplexed about her life. Nonetheless, in this paragraph I take as fundamental meaning in the autobiographical sense because it better takes into account a person’s authenticity – I can live not authentically and still have an existence judged as biographically meaningful, while the same thing would be difficult at the autobiographical level.⁴⁷

We now have to consider interactions between the already defined concepts. A happy existence will always be rationally worthwhile, but it can be insignificant (in both ways). A worthwhile life can be both happy and meaningful,

⁴⁶Metz [19] p. 447.

⁴⁷It could be objected that, without an external viewpoint, my position concedes that individuals find their own meaning in the most improbable or repugnant things, such as the search for flying donkeys or genocide – much as lives dedicated to not less problematic ends are admired by entire crowds. Actually, my claim about autobiographical meaning is based precisely on the refusal of extreme subjectivism in this field. It is a fact, moreover, that almost all “those currently philosophising about meaning in life would deem beneficence and creativity to confer meaning on a life, regardless of theoretical commitments in the background”(Metz [18] p. 8). I am not interested here in developing an original, objective notion of meaning, but simply in devising a generic, objective view based on reasonable assumptions.

and to remain so it must probably fulfill at least one of these conditions (intending meaning as autobiographical). An autobiographically meaningful life can be happy or unhappy, but it must be worthwhile. This last claim could seem counterintuitive. Indeed, Thaddeus Metz wrote:

Consider, say, someone who volunteers to be head of department, taking on administrative burden and attending dull meetings so that his colleagues can avoid doing so. Or, less glibly, think about those in the caring professions, such as nurses who elect to face stench, filth, distress, and the like so that such conditions are lessened for others. It is natural to say that such actions make these people's lives more meaningful, albeit not, or at least to a much lesser degree, worthwhile. Next, reflect on individuals who commit suicide for a good cause, such as protecting innocents. Take a classic lifeboat scenario where there are not enough seats for all those who need them, and where you volunteer to give yours to someone else – a meaningful action, albeit not one that would make one's life worth continuing.⁴⁸

Now, the first two examples only demonstrate that some highly meaningful activities can be sources of unhappiness – nonetheless, as I have already stressed, an unhappy existence could be very significant (in both ways) and then worthwhile anyway. The last example, on the other hand, seems to directly question my assumption that an autobiographically meaningful life will always be at the same time worthwhile: in the lifeboat case the choice of death is notably significant. However, this example is not conclusive.

I said that one of the ways in which an existence can become meaningful is when it relates to a context transcending the limits of its subject – e.g. an individual could devote her life to the promotion of peace between human beings. As the source of meaning is the relationship with the exceeding dimension, it is clear that, at least sometimes, the safeguard of such a dimension acquires a greater value than (the protection of) a single life dedicated to it. We could say that, in the case of a person living for the advancement of peace, while the latter is intrinsically significant, her existence is only extrinsically meaningful – in the autobiographical acceptance of the term. Imagine then that the person choosing to commit suicide by diving into the sea is a pacifist, and that the one to whom she gives her seat is the only person who, once arrived on land, will be able to avoid a global war. So, let us consider the pacifist's

⁴⁸Metz [19] p. 444.

alternatives: (a) staying on the lifeboat and repudiating her values, living a life almost certainly lacking deep (at least autobiographical) significance, and perhaps not even worthwhile (she will probably be tormented by the feel that she has done a wrong saving herself instead of trying to avoid a world war); committing suicide thus making her existence highly meaningful (moreover, from the very moment of the decision to the occurring of death she will be conscious to have done the right thing). Now, by choosing (b) our pacifist *is not* ceasing to judge her life worthwhile – if, before drowning, another, half-occupied lifeboat approached offering her a seat, she would obviously accept. She simply acknowledges that (1) the peacemaker’s life will be more meaningful and worthwhile than hers; (2) in her case, the promotion of peace between nation matters more than the prosecution of her life. In short, the pacifist comprehends that there is no place for both of them on the lifeboat.⁴⁹

Someone may object that our analysis of Metz’s lifeboat example, though correct, is not enough to exclude the possibility of formulating different, more compelling counterexamples to the thesis according to which an autobiographically meaningful life is always worthwhile. So I will now consider a more difficult case.

Take a war surgeon who decides to devote her life to save strangers’ lives. At a certain point, unable to bear the sensation of helplessness deriving from working in close contact with suffering and death, she deliberates to kill herself, not judging her existence worthwhile, though significant. The point, here, is that the word “significant” is not used in its autobiographical meaning. Without having to affirm that our surgeon comes to the point of denying the usefulness – and, then, probably also the meaning – of surgery in theatres of war, we can once again make reference to the notion of a meaningful life as a relationship with a wider context. This relationship is composed of two elements: in the current case, the surgeon’s existence and the importance of saving the lives of other people, irrespective of their membership group. When the surgeon feels that she cannot work anymore, she implicitly recognises the break between the two elements of the relationship, now incompatible: it will be impossible for her to operate once more, due to the too many horrors she has seen. The disruption of the ‘meaning-bond’ leads her to commit suicide because she thinks that her life cannot establish an equivalent bond with any other ele-

⁴⁹Note that we can say that killing oneself in order to rescue another person is a meaningful act only if we know the identities of the two individuals: if the pacifist had committed suicide by giving her seat to Adolf Hitler just before the invasion of Poland, her choice would not have been significant, but only stupid.

ment. Instead of living long, insignificant years she chooses to die. Then, if from a biographical standpoint her colleagues could suppose that she would have been able to continue working – meaningfully – for years, the autobiographical perspective suggests that the source of *her* meaning has been used up.⁵⁰

I think that it is now reasonable to affirm that, until proven otherwise, an autobiographically meaningful life is likely to be always worthwhile (even if sometimes barely, perhaps). After these clarifications, I will show their implications for Benatar’s anti-natalism.

In my opinion, the concept of autobiographical meaning (of a life) can demonstrate the wrongness of Benatar’s Basic Asymmetry; this is so because meaning is conceptually asymmetric – but in a different sense than that used by Benatar. Significance has no contrary. It cannot be inserted in the proportion $a : b = c : d$, i.e. benefit : harm = pleasure : pain, because nothing is for meaning what pain is for pleasure. Obviously, the absence of meaning is a negative thing, but not as the affirmation of something intrinsically bad (e.g. suffering). Instead, it is negative only as the absence of something intrinsically good. Furthermore, a meaningless life can still be happy and worthwhile, as in the classic example of the hedonist.⁵¹ A first clue of the fact that meaning (of life) lacks a contrary comes from language: “[t]here is apparently no English word denoting the negative correlate of a meaningful life”.⁵² Nonetheless, someone may consider the elaboration of an opposite concept possible:

Lives that are worthy of great esteem identify with something greater, and connect with goods worthy of love and allegiance have correlates: lives that are worthy of great shame, identify with something lesser, and connect with bads worthy of hatred and opposition. [...] Many are inclined to judge meaningless live to be downright bad, not merely lacking in goodness.⁵³

However, this approach has to overtake a notable obstacle: usually, people who commit acts judged highly insignificant (such as many murders) believe

⁵⁰Here I do not want to argue that in all the possible worlds the surgeon will never find another notable source of meaning. I am only putting in evidence that her decision can be authentic and rational given the elements she knows.

⁵¹Metz [19] p. 443.

⁵²Metz [17] p. 806.

⁵³Ibid.

that their conduct was just, or at least justified as a lesser evil. Take crimes of passion or crimes of honour: their architects often commit them thinking to atone for an injustice by private means. Only a mentally ill person could kill someone because of a generic hate against all human beings, or steal in order to promote dishonesty as a value.⁵⁴ A similar phenomenon regards people devoted to meaningless non-criminal activities, as in the case of a hypothetical flying donkeys searcher.

It could be objected that the autobiographical illusions of the agents are of no importance here: only objective, biographical meaning of one's actions is fundamental in weighting her life. This is a problematic conclusion: intentionality is a constitutive feature of the meaning of an act. Consider, for example, Metz's aforementioned definition of significance, which makes reference not only to the consequences of our actions, but also to their causes. The reason underlying this consideration is easy to comprehend. Imagine a secret agent believing to work for Country X, which is her country, whereas she is unconsciously saying one thing and doing another for Country Y, which is X's worst enemy. Now, let us suppose that she succeeds in blocking the theft of Y's secret military project at the hand of X. Y's citizens will be glad about her conduct, and they will probably judge her action useful, but they will certainly not consider it meaningful, because of her completely opposite intentions. As the example shows, affirming that there are persons promoting a sort of anti-meaning with the same level of consciousness typical of people devoting their lives to meaning is very difficult. Someone may notice that not all real cases are well described by my example: consider a serial killer murdering her victims on commission. She perfectly knows that killing is wrong, the objection goes, nonetheless she does it. However, I would reply that she kills for money, not for pleasure – best: she kills in spite of displeasure. Her moral code tells her that it is somewhat licit to assassinate if one does it in order to earn a living. Still, in a serial killer's existence there are surely relationships she values not for materialistic convenience, but for the meaning she attaches to them. This explain why, while a mentally sane serial killer (of course, assuming that such a person could exist) needs an incentive from outside the dynamics of the murder in order to kill, she does not need a similar type of incentive, for example,

⁵⁴13 It is important to notice that, though compatible with it, my view does not imply the acceptance of moral intellectualism, which affirms that, when a person commits a morally wrong act, this always happens due to her ignorance of what is morally good. Indeed, in many cases people acting wrongly may know what would be right to do, but nonetheless lack the possibility or the determination sufficient to bring it about. However, failing to perform a morally right act could hardly be regarded as a deliberate promotion of what is wrong.

to love somebody.⁵⁵

Then, if my argument is sound, meaning cannot be taken as part of Benatar's scheme describing harms and benefits. Nonetheless, he could still claim that, though the presence of (biographical and, above all, autobiographical) significance is good, its absence is not bad in a scenario in which nobody would be deprived of it. That is, also a more precise account of this concept would not avoid anti-natalist implications. But if anti-natalism is based on the assumption that life is never worth-starting, an assumption partly relying on the very distortion of the structure of meaning of life I have highlighted (and also caused by a logically dubious conclusion, as well as by an incorrect reading of psychological dynamics), I am driven to think that it provides a quite indefensible view.

There is still a further reason in favour of this last point: given that having children is usually meaningful for a person, why should she avoid it, if their existences are likely to be significant too? As an implicit demonstration of the importance of meaning when judging an existence worthwhile, I will now examine Benatar's interpretation of suicide:

Potential procreators would do well to consider this trap they lay when they produce offspring. It is not the case that one can create new people on the assumption that if they are not pleased to have come into existence, they can simply kill themselves. [...] One's life may be bad, but one must consider what affect ending it would have on one's family and friends.⁵⁶

Now, the fact that many persons, though unhappy, are unable to commit suicide because of the consequences that it would have on their loved ones, simply demonstrates the ability of meaningful relationships to make a sad existence worthwhile. Some people seem to limit our freedom because we care about them. The impossibility of killing oneself under some circumstances has not to be explained in terms of (logical or causal) necessity – indeed, we can make the opposite decision. In these cases we do not lack power, we lack will: it is a matter of that “volitional necessity” coinciding with desires that are not exactly ours, but with which we actively identify ourselves.⁵⁷ Obviously, as

⁵⁵Clearly, this is not only due to the fact that killing is forbidden by the law: many people would (at least try to) continue to have meaningful love relationships even if they were as illegal as murder.

⁵⁶Benatar [2] p. 220.

⁵⁷Frankfurt [8] p. 264–265.

Benatar notes, this is not the whole story: some people suffer so much that their loved ones cannot ask them continuing to live unless they are egoists.⁵⁸ However, such cases are extreme.

A final objection must be mentioned here: from the perspective of the universe, a human life could hardly be meaningful at all – and autobiographical meaning is probably part of the wider phenomenon of positive illusions.⁵⁹ Furthermore, according to Benatar, judging lives *sub specie aeternitatis* is, for example, the best way to explain why modesty is a virtue. But his argument against meaning is implicitly founded on several problematic assumptions. First, if someone endorses an objective theory of the meaning of life, she could have no problem in recognising that what is significant in her existence can also be meaningful in a more general perspective. Then, Benatar seems to reject a theory of this kind without discussion. Second, the very idea of evaluating a life *sub specie aeternitatis* is problematic: eternity and the universe are not subjects, and so they cannot have a point of view. For this reason, when Benatar suggests looking at one's life from their angle, it is not clear what he has in mind. Third, many theories of (objective) meaning actually adopt an eternal standpoint: God's one. Nonetheless, they judge human life meaningful precisely because such a Being exists. It is quite evident that Benatar does not believe in God, but if he wants to formulate hypotheses *sub specie aeternitatis*, he has to say something about the reason why we should exclude *a priori* a theistic version of that viewpoint (more on this in the next paragraph). Fourth, if no life is really meaningful, then (human) ethics itself may be meaningless. Let us imagine for a second that we accept Benatar's vision of coming into existence as always a harm – and then of procreation as something morally problematic. How important would the perpetuation of this harm be from the perspective of the universe? It seems to be a very controversial question. So, I am not sure that Benatar can deny the importance of meaning in one's existence without compromising the role of ethics itself – and, even if my worry was exaggerated, he would still have to demonstrate why. Fifth, we have already seen that Benatar's account of positive illusions is weak. Finally, I do not think that a person is modest when, although her accurate perception of her strengths, she recognises that there is a *sub specie aeternitatis* standard by which she falls short.⁶⁰ In my opinion, modest people are simply aware that put aside their strengths, (a) there are a lot of important

⁵⁸Benatar [2].

⁵⁹Ibid. pp. 82–86.

⁶⁰Ibid. p. 86.

contexts in which they are not so good; (b) it is wrong to talk big about one's qualities and results, because it could make other people feel bad.

6 Anti-natalism and Religious Views: an impossible coexistence

Before formulating an alternative explanation of Benatar's four other asymmetries, in this paragraph I will briefly analyse his attempt to conciliate anti-natalism with religion in general – and Christianity in particular. In order to find biblical support for his arguments, he makes reference to three passages from the Old Testament: Ecclesiastes 1, 2–4; Jeremiah 20, 14–18; Job 3, 2–4, 6, 10–11, 13, 16. In all these excerpts, the protagonist declares that he would prefer never to have been. Benatar interprets these extracts as the demonstration that a deep and non-fundamentalist attitude toward religion could incorporate his core ideas.⁶¹

Unfortunately, his biblical exegesis is untenable. Indeed, for all the biblical figures Benatar mentions it cannot be said, as he seems to do,⁶² that pleasure and pain (better: benefits and harms) are the salient parameters when making a distinction between a life which is worthwhile and one that is not: for them the only relevant variable is their relationship with God.⁶³ More precisely, it is when He appears insensible to the injustices happening in the world that their judgment on life becomes negative.⁶⁴ Furthermore, in a Christian reading of the Old Testament, the impossibility of redemption in Qohelet's nihilism or in Job's desperation appears radically overcome.⁶⁵ Similarly, in Kierkegaard's view Job is the man who, willing to dispute with his Lord, definitively detaches himself from the old retributive logic, and builds a bridge between Judaism and Christianity – in which the proportional correlation among guilt and punishment is replaced by the overabundant divine mercy.⁶⁶ We can then understand that the positive (at least partially) epilogue of Job's book is no coincidence.

I have tried, with these few remarks, to highlight the fact that Benatar's in-

⁶¹Benatar [2] pp. 222–223.

⁶²Ibid. p. 22.

⁶³Curi [6] p. 193.

⁶⁴Ibid. pp. 210; 212–213; 230.

⁶⁵Ibid. pp. 207; 218.

⁶⁶Ibid. pp. 239–245.

terpretation of the Bible could be accepted at best by non-Christians – thus lacking any argumentative power. However, there is an even more radical reason why Benatar’s anti-natalism can be interpreted as a negation of any kind of transcendent dimension – and so of religion. His evaluation of the quality of human existence does not admit the possibility of an afterlife. From Benatar’s perspective, affirming that a life is not worth-starting implies that it ends with death. Now, according to many philosophies of religion, biological existence is a necessary prerequisite for eternal life and procreation has the status of a (moderate) duty, so Benatar implicitly asks us to bet, against Pascal, on God’s absence. This is a very risky request, also considering that even amongst non-religious philosophers anti-natalism is significantly unpopular. In the end, Benatar wants us to end human life for motives incompatible with religion and strongly criticised also by atheist philosophers: it seems to me to be an unreasonable demand.

7 The Four Asymmetries: an alternative analysis

In this paragraph, I will provide a critical analysis of Benatar’s other asymmetries, demonstrating that, even if we accept all of them, they can be explained without making anti-natalistic claims.

The asymmetry of procreational duties. While we have a duty to avoid bringing into existence people who would lead miserable lives, we have no duty to bring into existence those who live happy lives.⁶⁷

I believe that this asymmetry can be explained by making reference to the notion of possibility: whereas there are phenomena (e.g. serious malformations, life-threatening diseases presently incurable, unsuitable environments and conditions for human life) which can make us really sure of the fact that an existence would not be worthwhile, there are not, on the contrary, elements sufficient to make the statement that a life will surely be worth-living – consider the story of Croesus and Solon narrated by Herodotus. Because phenomena of the first kind are often foreseeable and not so widespread in many coun-

⁶⁷Benatar [3] p. 123. I think that there is at least a possible case in which this asymmetry does not hold (cf. Narveson [21] p. 72). However, in what follows I assume, for the sake of argument, that the procreational duties asymmetry is always true.

tries,⁶⁸ it is reasonable to say that lives are worth-starting in numerous cases.⁶⁹ This implies that procreation cannot, at least under normal circumstances (see note 67), be described as a duty. A human existence will always contain an endemic part of risk for its subject: if we sometimes are morally obliged to personally take a chance, only in extreme situations we can consider the possibility of a duty to make others run a risk. Then, if not to come into existence represents the (ontological) absence of possibility, starting to exist exemplifies the intrinsic ambivalence of the category itself: in a Kierkegaardian way, every possibility of achievement is, at the same time, a possibility of failure. In discussing of possible future people this uncertainty is particularly evident because even happy people cannot be sure of the overall worthwhileness of their lives – as the epilogue in Croesus' story suggests.⁷⁰

Only after having noticed such an ontological ambivalence, we are ready to comprehend the exact meaning of the ancient maxim according to which the best thing is not to be born, while the second best is to die soon after. It is not a phrase that can be interpreted to the letter (i.e. as a form of anti-natalism), but an expression of oracular wisdom that recognises our impossibility to know what is better for us.⁷¹

In conclusion, understanding the importance of the concept of possibility also allows us, on a prudential basis, not to create those existences whose unworthiness is not completely sure, but probable (as this will somewhat negatively polarise the ambivalence previously mentioned).

The prospective beneficence asymmetry. It is strange to cite as a reason for having a child that that child will thereby be benefited. It is not similarly strange to cite as a reason for not having a child that that child will suffer.⁷²

⁶⁸If we did not systemically violate our duties towards the world's poor, this claim could probably be made for the entire world (on this topic see, for example, Pogge [24]).

⁶⁹Benatar [2] clearly distinguishes between lives worth-starting and worth-continuing (pp. 22–28). Because this distinction is founded upon an argument that I have already rejected when discussing his view of suicide, I will not consider it in more detail. An additional, thoughtful objection can be found in DeGrazia [7] (p. 321).

⁷⁰However, it is probable that the sense of insecurity would be less when thinking of a life's meaning.

⁷¹Curi [6] pp. 70–72; 279–283.

⁷²Benatar [3] p. 123.

Taking in mind the explanation of the previous asymmetry, it is easy to comprehend this one: if, in some cases, coming to existence can be a final sentence to the absence of worthwhileness (or happiness), it can never be a final sentence to worthwhileness (or happiness). This simple fact elucidates why two healthy carriers of the genes of a terrible, incurable disease should choose not to have children: it is rational to say that the possible baby of the couple would have a very unhappy and insignificant existence. However, it would not be rational to affirm, with the same certainty, that the possible offspring of two perfectly healthy parents will have a worthwhile life. (Someone could claim that this demonstrates the wider asymmetry between good and bad which Benatar is – perhaps –looking for, but this would simply be false: the obvious asymmetry, here, is between existing and not existing).

The retrospective beneficence asymmetry. When one has brought a suffering child into existence, it makes sense to regret having brought that child into existence – and to regret it for the sake of that child. By contrast, when one fails to bring a happy child into existence, one cannot regret the failure for the sake of the person.⁷³

It seems evident to me that what matters here is the (ontological) weight of an existing, unhappy child, against the lesser importance of the unhappiness of a non-existing person, as suggested by David DeGrazia:

[R.]ather than explaining this conjunction of claims in terms of a fundamental asymmetry between harm and benefit, I explain it by reference to the fact that in the former case there is a subject who can be harmed, whereas in the latter case there is no subject to be harmed. On the other hand, if the conjunction of claims is best understood in impersonal terms, then I see no reason to accept the present asymmetry. For, if it is impersonally bad for someone with a predictably bad life to come into being, it seems that it would be impersonally good for someone with a predictably good life to come into being.⁷⁴

⁷³Ibid.

⁷⁴DeGrazia [7] p. 323. N.B. Benatar objected that these claims, combined with the view, also defended by DeGrazia, that we have stronger duties not to harm than to benefit, yields the anti-natalist conclusion via a different route (Benatar [3] pp. 131–132). I think that this objection is weakened by my refutation of the Second Argument but, in any case, because I do not endorse DeGrazia's view on the asymmetry between negative and positive duties, the coherence of my own position is out of danger.

This is, once again, a consequence of the intrinsic structure of possibility (roughly speaking: there are many ways of existing, but only one of non-existing).

Before analysing Benatar's last asymmetry, I will very briefly consider a possible criticism against the view I have developed in this paragraph. If we cannot be sure that the lives we create will be worth living – the objection goes – we should, on a (nihilistically) precautionary basis, stop creating lives at all. Now, given the fact that – at least under some circumstances – the great majority of the existences we create are worthwhile, criticism of this kind is unacceptable because it would lead to a variant of what Derek Parfit called “The Ridiculous Conclusion”:⁷⁵ it would attribute to harms an exponentially higher consideration than that attributed to benefits.

Now, let us turn to the fourth asymmetry:

The asymmetry of distant suffering and absent happy people. We are rightly sad for distant people who suffer. By contrast we need not shed any tears for absent happy people on uninhabited planets, or uninhabited islands or other regions on our own planet.⁷⁶

As I anticipated, I do not consider this a real asymmetry: Benatar compares regret for real suffering people and the lack of compunction for the happiness of nonexistent possible persons, but this says nothing about an asymmetrical relationship between benefits and harms. Imagine the opposite asymmetry, that in which we put contentment for happy real people alongside the absence of happiness for possible sad persons who luckily do not exist: it seems as intuitively valid as its companion. The problem is that two parallel asymmetries are logically absurd: they would form a new symmetry.

Benatar may object that, once accepted his retrospective beneficence asymmetry, we must concede that the nonexistence of possible, unhappy people should make us happy, but this argument would not be exact. The third asymmetry only affirms that we would feel guilty with X if, making her exist, we condemn her to a life which is not worthwhile. But this last set of propositions does neither imply (1) that not making X exist is a source of happiness for us, nor (2) that thinking of the possible Xs who would live, if existed, a not worth-living existence on a deserted island, would make us happy. For these reasons I am of the opinion that it would be better for Benatar to give

⁷⁵ See Parfit [23] pp. 401–412.

⁷⁶ Benatar [3] p. 123.

away his fourth asymmetry, otherwise he will be forced to accept the logically problematic consequences of its opposite.

Finally, even if we accepted the last asymmetry, there would be at least two other possible explanations for its content. The first relies on DeGrazia's objection before mentioned, affirming that nobody is personally harmed by the lack of happiness in uninhabited regions of the planet and so we have nothing to regret it. The second, on the other end, affirms that we do not regret the absent pleasures of those who could have existed simply because anti-natalism is not widespread. In other words: it is possible to believe that the global population has approximately reached a threshold crossing which its rapid growth would disadvantage already existing people more than it would advantage those who could exist. Alternatively: our planet is characterised by a population density sufficiently high in many areas that it can offer us only a small chance of regretting the absence of people in unoccupied regions.

8 Conclusion

In this article I have critically analysed David Benatar's arguments in favour of anti-natalism, reported in paragraph 2 and later at several stages in the text. I firstly tried to show the initial weakness of Benatar's Basic Asymmetry (paragraph 3), focusing then on the improbability of the psychological assumptions of his Second Argument, based on an unacceptable kind of pessimism (paragraph 4). In paragraph 5 I developed an original account of meaning in life, claiming that the notion of meaning is asymmetric in a way that constitutes in itself an original objection to anti-natalism (and also to Benatar's view of suicide). In paragraph 6, I briefly noted that Benatar's core propositions are incompatible with religion – and in particular with Christianity. Paragraph 7 provided a single, different, and clear explanation of other three asymmetries mentioned by Benatar in support of his Basic one, offering both a rejection and a set of alternative explanations for a fourth asymmetry.

In conclusion, it seems to me that there is enough evidence to reject at all Benatar's claims. I have to acknowledge, however, that this conclusion is due to the philosophical unsustainability of anti-natalism *per se*, rather than because of that of Benatar's own proposal – which, although mistaken, remains extraordinarily original and thoughtful.

References

- [1] Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001) 'Bad Is Stronger Than Good', in *Review of General Psychology*, 5(4): 323–370.
- [2] Benatar, D. (2006) *Better Never to Have Been: The Harm of Coming into Existence*. Oxford: Oxford University Press.
- [3] Bentar, D. (2013) 'Still Better Never to Have Been: A Reply to (More of) My Critics', in *The Journal of Ethics*, 17: 121–151.
- [4] Brickman, P., Coates, D., & Janoff-Bulman, R. (1978) 'Lottery winners and accidents victims: Is happiness relative?', in *Journal of Personality and Social Psychology*, 36: 917–927.
- [5] Bruckner, D. (2009) 'In defense of adaptive preferences', in *Philosophical Studies* 142: 307–324.
- [6] Curi, U. (2008) *Meglio non essere nati: La condizione umana tra Eschilo e Nietzsche*. Turin: Bollati Boringhieri.
- [7] DeGrazia, D. (2010) 'Is it wrong to impose the harms of human life? A reply to Benatar', in *Theoretical Medicine and Bioethics* 31: 317–331.
- [8] Frankfurt, H. (1982) 'The importance of what we care about', in *Synthese*, 53: 257–272.
- [9] Hare, R. (1993) *Essays on Bioethics*. Oxford: Oxford University Press.
- [10] Harman, E. (2009) 'Critical study of Benatar (2006)', in *Noûs* 43: 776–785.
- [11] Haybron, D. (2011) 'Happiness', in *Stanford Encyclopedia of Philosophy*. [Online] Available at: <http://plato.stanford.edu/entries/happiness/>
- [12] Huemer, M. (2008) 'In Defense of Repugnance', in *Mind*, 468: 899–933.
- [13] Hunston, T.L. & Vangelisti, A.L. (1991) 'Socioemotional behavior and satisfaction in marital relationships: A longitudinal study', in *Journal of Personality and Social Psychology*, 61: 721–733.
- [14] Kahneman, D. & Krueger, A.B. (2006) 'Developments in the Measurement of Subjective Well-Being', in *Journal of Economic Perspectives*, 20: 3–24.

- [15] Levy, A. (2005) ‘Liberalism, Adaptive Preferences, and Gender Equality’, in *Hypatia* 20(4): 127–143.
- [16] McMahan, J. (2009) ‘Asymmetries in the Morality of Causing People to Exist’, in M.A. Roberts & D.T. Wasserman (ed.), *Harming Future Persons*. Dordrecht: Springer.
- [17] Metz, T. (2002) ‘Recent Work on the Meaning of Life’, in *Ethics*, 112: 781–814.
- [18] Metz, T. (2009) ‘Happiness and Meaningfulness: Some Key Differences’, in L. Bortolotti (ed.), *Philosophy and Happiness*. Houndmills: Palgrave Macmillan.
- [19] Metz, T. (2012) ‘The meaningful and the worthwhile: clarifying the relationships’, in *The Philosophical Forum*, 43: 435–448.
- [20] Nahmias, E., Morris, S., Nadelhoffer, T. & Turner, J. (2005) ‘Surveying Freedom: Folk Intuitions about Free Will and Moral Responsibility’, in *Philosophical Psychology*, 18: 561–584.
- [21] Narveson, I. (1967) ‘Utilitarianism and New Generations’, in *Mind*, 76(301): 62–72.
- [22] Nussbaum, M. (2001) ‘Adaptive preferences and women’s options’, in *Economics and Philosophy*, 17: 67–88.
- [23] Parfit, D. (1987) *Reasons and Persons*. Oxford: Oxford University Press.
- [24] Pogge, T. (2008) *World Poverty and Human Rights: Cosmopolitan Responsibilities and Reforms*. Cambridge: Polity Press.
- [25] Ritchie, T., Skowronski, J.J., Hartnett, J., Wells, B. & Walker, W.R. (2009) ‘The fading affect bias in the context of emotion activation level, mood, and personal theories of emotion change’, in *Memory*, 17: 428–444.
- [26] Sheldon, K.M., Ryan, R., & Reis, H.T., (1996) ‘What makes for a good day? Competence and autonomy in the day and in the person’, in *Personality and Social Psychology Bulletin*, 22: 1270–1279.
- [27] Taylor, S. (1991) ‘Asymmetrical Effects of Positive and Negative Events: The Mobilization-Minimization Hypothesis’, in *Psychological Bulletin*, 110: 67–85.
- [28] Taylor, S. & Armor, D. (1996) ‘Positive Illusions and Copying with Adversity’, in *Journal of Personality*, 64: 873–898.

- [29] Walker, W.R., Vog, R.J., & Thompson, C.P. (1997) 'Autobiographical memory: unpleasantness fades faster than pleasantness over time', in *Applied Cognitive Psychology*, 11: 399–413.
- [30] Walker, W.R. & Skowronski, J.J. (2009) 'The Fading Affect Bias: But What the Hell Is It For?', in *Applied Cognitive Psychology*, 23: 1122–1136.

How Convincing is the Levelling Down Objection to Egalitarianism?

Joel Griffith

University of Exeter

The Levelling Down Objection (LDO) is perhaps the most prevalent and powerful anti-egalitarian criticism that exists today. Its proponents aim to deflate the egalitarian position by arguing that, sometimes, creating a situation that is ‘equal’ improves the well-being of nobody, and some, namely those better off, are harmed. In this essay, I argue that the LDO can be rejected on numerous grounds. I shall begin with a clarification of how I believe the concept of egalitarianism should be interpreted; following the method of Larry Temkin, I shall focus in particular on the frequent conflation between prioritarianism and egalitarianism. I’ll then proceed to suggest that much of the LDO’s force is derived from a person-affecting view called the Slogan.¹ An assessment of the Slogan and its implications exposes three significant shortcomings concerning (1) Derek Parfit’s non-identity problem, (2) proportional justice, and (3) what I shall call ‘the problem of suffering’. Next, I argue that there are two variants of egalitarianism not considered by Parfit that evade the LDO. Firstly, a position that holds equality as extrinsically but non-instrumentally valuable, and, secondly, a position that holds equality as intrinsically and non-instrumentally valuable. As we shall see, these two approaches seem to be person-affecting forms of egalitarianism. However, a consideration of the distinction between pure and pluralist (O’Neill [4]) egalitarian standpoints can be used to successfully defend these approaches. Ultimately, I wish to conclude that the LDO itself does little to undermine the plausibility of egalitarian claims.

Many who label themselves egalitarians subscribe to a view like the following: they want each person to fare as well as they possibly can; however, they are especially concerned with those who are worse off.² Thus, in any event, the worse off someone is, the greater priority they receive in our moral deliberations, particularly concerning redistribution. Derek Parfit has called such a view ‘The Priority View’;³ however, for simplicity, I shall follow Larry

¹Temkin [6].

²Ibid.

³Parfit [5]

Temkin and refer to it as ‘prioritarianism’.

As a strand of egalitarianism, prioritarianism faces a number of problems. Consider Figure 1, where the column heights represent how well off people are, and the widths represent the number of people in each group.

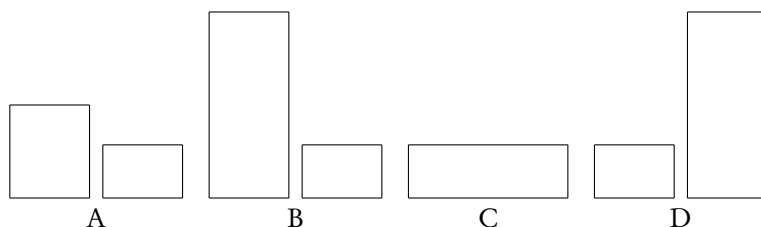


Figure 1

According to prioritarianism, there would be no reason for one to prefer A over B, despite it's obviously being more equal. Likewise with C and D respectively. As an egalitarian position, the issue with prioritarianism is clear. It is not concerned with equality *per se*.⁴ Equality describes a relation that is necessarily *comparative*. People are more or less equal relative to one another. In contrast, prioritarianism is concerned with how people fare, but not with how they fare relative to each other.

Since many prioritarrians think of themselves as egalitarians, it may be useful at this point to distinguish between what Derek Parfit describes as deontic and telic egalitarianism.⁵ According to deontic egalitarianism, equality is extrinsically valuable, i.e. valuable when it promotes some other ideal, such as well being or utility. In contrast, telic egalitarians believe that equality is intrinsically valuable – that is, valuable as an end in itself, rather as a means by which we can promote other ideals.⁶ Crucially, as we will see later when examining pluralist perspectives, telic egalitarians need not believe that equality is all that matters – they hold that equality is one ideal, among others, that possesses independent moral value.

To sum up, prioritarianism is often conflated with egalitarianism. As we have seen, prioritarianism expresses a special concern for the worse off, yet it is

⁴Temkin [6].

⁵Parfit [5].

⁶Ibid.

not concerned with how the worse off fare relative to others. As a result, as Figure 1 illustrates, it licenses vast increases in inequality in the name of improving conditions for those that are worse off. Of course, giving priority to such individuals will generally promote equality through a system of redistribution. Hence, we can regard prioritarianism as a plausible deontic egalitarian position. However, in this respect, prioritarianism is not unlike utilitarianism, which also favours benefiting the worse off as a way of increasing utility; neither are plausible as a telic egalitarian position, since neither values equality *per se*.

Temkin suggests that people are drawn to prioritarianism since it is the closest thing to an egalitarian position one can plausibly adopt. The underlying premise here is not that prioritarianism is a plausible form of telic egalitarianism, but rather that telic egalitarianism is implausible.⁷ Many are attracted to the foregoing by the LDO – Figure 2 illustrates this.

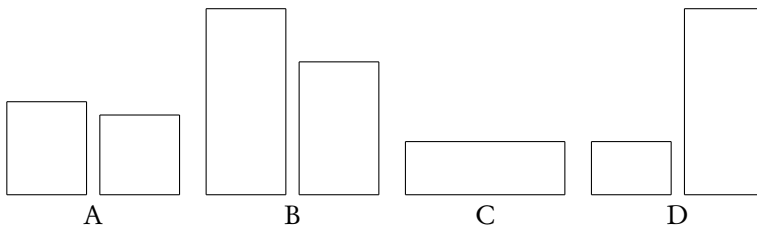


Figure 2

Imagine we could switch from A to B. Many find it hard to believe that there could be any reason not to do this. In B, every individual is better off than they were in A. Even B's worse off have better lives than A's better off. True, there is greater inequality in B than A; however, doesn't this indicate that we shouldn't attach weight to equality *per se*? How can the inequality in B be bad, when there is nobody for whom it is worse in comparison to A?⁸

It is clear from this why such considerations have been dubbed 'the LDO'. Telic egalitarianism, as we have seen, attaches value to equality *per se*. So, a telic egalitarian would in fact support a switch from B to A, rather than from A to B, by levelling down the relevant groups. However, such a move benefits nobody, not even the worse off. In fact, a move from B to A would

⁷Temkin [6].

⁸Ibid.

significantly harm the worse off. In such cases, many argue that surely there is nothing that can be said in favour of promoting greater equality, which is desirable only when it benefits the worse off, not when it results from levelling down. Hence, the LDO concludes that equality is *extrinsically* valuable and therefore, telic egalitarianism should be rejected.⁹

At the heart of the LDO is a view described by Derek Parfit as the ‘person-affecting claim’, and by Larry Temkin as ‘the Slogan’, which reads: ‘One situation cannot be worse (or better) than another if there is no-one for whom it is worse (or better).’¹⁰ Outcomes should be assessed solely in terms of the way the beings in those outcomes are affected, for better or worse. I believe it is the Slogan that gives the LDO much of its rhetorical force. If we reject it, then there seems to be little basis for rejecting telic egalitarianism as a result of the LDO. However, despite being seemingly invoked in many cases, the Slogan can, and should be, challenged.

It is without doubt that the Slogan has great force and initial appeal. Nevertheless, it must be rejected. Firstly, consider a variation of Parfit’s non-identity problem, illustrated below in Figure 3.

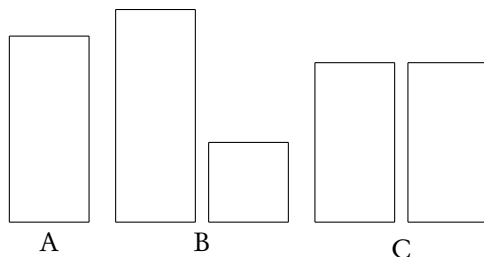


Figure 3

Let A represent a generation contemplating two policies. On the ‘live for today’ policy, they have children immediately and deplete natural resources for current use; B would result. On the ‘take care of tomorrow’ policy, they postpone having children and conserve resources; C would result. Most believe that the latter policy should be adopted, yet this cannot be accounted for by the Slogan. This is because of the non-identity problem. Firstly, the children born in C would be identifiably different to those in B (they would be con-

⁹Parfit [5].

¹⁰Temkin [6].

ceived later). Secondly, one cannot act against the interests of someone who will never exist. Given these two points, there is no one the live for today policy affects for the worse: not the parents, who fare better in B than in either A or C; not the children in B, because they wouldn't exist if the take care of tomorrow policy was adopted; and not the children in C, because they don't exist and never will exist if the live for today policy is adopted.¹¹ On the other hand, if the take care of tomorrow policy is adopted, there will be someone adversely affected, namely the parents. According to the Slogan, then, the live for today policy cannot be worse than the take care of tomorrow policy, since there is no one for whom it is worse. But this is surely wrong. The live for today policy is worse than the take care of tomorrow policy. Thus, the Slogan must be rejected or at least is limited in scope.

Many have argued that the above problem shows that there is merely a limited and peculiar range of cases where the Slogan does not apply – particularly when our choices determine who comes to be. However, this cannot be the case. Consider Figure 4, which I shall use to illustrate the concept of proportional justice.

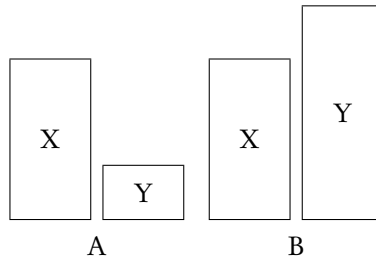


Figure 4

Let A and B represent alternative afterlives, with X representing good people, and Y bad people. Also, assume A reflects how the two groups *should* fare according to a system of justice (i.e. a system in which people get what they are worth/deserve). Clearly, in this instance, A would be better than B. However, according to the Slogan, the reverse would be true. In fact, there would be no respect in which B would be worse than A. This is particularly hard to accept – surely there is something morally inadequate about bad people faring better than good people, even if there is nobody for whom it is worse. These considerations suggest that unless one is willing to reject proportional justice entirely,

¹¹Temkin [6].

one must reject the Slogan. Obviously, the former seems greatly more appealing. The principle of proportional justice is most plausibly interpreted as an impersonal principle. It assesses outcomes according to what people deserve rather than merely in terms of whether people are affected for better or worse. Thus, it is not surprising that the principle of proportional justice conflicts with person-affecting claims. As soon as one attributes value to something beyond its ability to affect beings for better or worse, one has carved out a role for impersonal principles in the assessment of outcomes, which in turns leads to a rejection of any sort of hegemonic person-affecting view like the Slogan.

The non-identity problem and issues concerning proportional justice make the Slogan look dubious at best. However, before I move on, let me point out another related problem that faces it – I shall call this ‘the problem of suffering’.¹² Imagine that individuals in society A have a life so bad that it is worse than nothing, i.e. not being alive. In B, there are no people. Since people suffer so much in A, it seems plausible to suggest that B would be the preferable option. However, according to the Slogan, B cannot be better, since it is better for no-one. This example illustrates the limitations of the Slogan in terms of practical implication. It only works when there is a comparative element to the situation at hand, i.e. at least two groups to which we can attribute a classification of how well off they are. Thus, like in the aforementioned example, when we have to consider a situation in abstraction from the normal person-affecting relationships, i.e. a system in which nobody exists, the Slogan fails to provide enough versatility to provide an adequate answer.¹³

So far, I have explored the credentials of the LDO within the framework of Parfit’s distinction between deontic and telic egalitarianism. To recap briefly, telic egalitarians accept what Parfit calls the principle of equality, namely that it is bad in itself if some people are worse off than others.¹⁴ Deontic egalitarians, in contrast, do not endorse this principle. When they object to inequality, they do so on the grounds that it involves some extrinsic wrongdoing, not because it is intrinsically bad. This distinction is of crucial importance, since telic egalitarianism is vulnerable to the LDO whilst deontic egalitarianism can

¹²Holtug [1].

¹³For an interesting insight into a possible revision of the Slogan, see N. Holtug’s paper ‘Egalitarianism and the Levelling Down Objection’ – though I shall not focus on his assertions in this paper, since I believe he offers a somewhat *ad hoc* adjustment.

¹⁴Parfit [5].

avoid it all together.¹⁵ Deontic egalitarians can escape this objection because they maintain that “we have a reason to remove inequality only when and only because, our way of doing so benefits people who are worse off”.¹⁶

This distinction is illuminating. However, I believe Parfit fails to consider two possible forms of egalitarianism that can also avoid the LDO.¹⁷ In order to identify these positions, a distinction needs to be made between non-instrumental and intrinsic value. When something has non-instrumental value it has value for its own sake. When something has intrinsic value, the value it possesses is grounded entirely in its intrinsic properties which do not depend on the existence or nature of something else.¹⁸ In contrast, extrinsic value is not wholly grounded in intrinsic properties. Making these distinctions allows us to identify the possibility of two forms of egalitarianism neglected by Parfit.

The first of these positions holds that equality is extrinsically but non-instrumentally valuable. It is extrinsically valuable because it is valuable for its own sake only when at least some other people benefit from it. Whether these people benefit from equality is to be determined by considering other empirically possible states of affairs and drawing a subsequent comparison to deduce whether there is one in which everyone is better off than they would be under equality. According to this view, equality is non-instrumentally valuable if and only if there are no such states of affairs, i.e. a system in which people are better off than in equality. When individuals benefit from equality, we should value it for its own sake. Some might object that if equality’s value is derived from its ability to benefit some, then it is not valuable for its own sake, strictly speaking. However, there seems to be no pressure to insist that if the value of something is partly grounded in the meeting of another condition, then it cannot be non-instrumentally valuable.¹⁹

The second position holds equality as intrinsically and non-instrumentally valuable but only when it benefits at least some. When equality benefits nobody it lacks both intrinsic and non-instrumental value. This, at first, may seem incoherent: if equality is intrinsically valuable on some occasions, then

¹⁵Ibid.

¹⁶O’Neill [4].

¹⁷Mason [2].

¹⁸Ibid.

¹⁹Mason [2].

surely it must be intrinsically valuable whenever it obtains.²⁰ However, the intrinsic value of something is any value grounded in intrinsic properties. This does not rule out the possibility that something may be of intrinsic value in one context and not another, for such a value may be affected by the context in which it appears. Therefore, this second position maintains that equality is valuable for its own sake only in the presence of an enabling context, namely it's benefitting some, and when this is present, its value is grounded entirely on the intrinsic properties of equality.

It is unsurprising that neither of these positions fit well into Parfit's classification system, since, by definition, they cut across the distinction between deontic and telic egalitarianism. Both, I believe, manage to avoid the LDO, since the LDO relies on the claim that levelling down cannot constitute an improvement in any respect – they agree with this claim. Crucially, both these variants will require an examination of possible states of affairs in which people are better or worse off – thus they resemble person-affecting forms of egalitarianism.²¹ As we have seen, Larry Temkin argues that person-affecting views are not transitive. He gives the following example to illustrate this. From the standpoint of a person-affecting claim, an outcome in which all are sighted is better with regard to equality than an outcome in which some are sighted and some are blind.²² However, it does not follow that an outcome in which all are sighted is better regarding equality than an outcome in which all are blind – these are equivalent regarding equality – since there are no considerations made in abstraction from merely the better/worse off relation of the individuals involved.

The versions of egalitarianism I have distinguished from Parfit's classification system may be regarded as counter intuitive in some respects. They both regard equality as non-instrumentally valuable even when the alternative is an inequality that would make all but one individual massively better off. Here, Parfit's distinction between pure and pluralist egalitarianism can do some work.²³ Pure egalitarians care only about equality, whereas pluralist egalitarians care about both equality and utility. Thus, the latter can regard a state of affairs in which equality obtains as worse than one where there is inequality but all but one are massively better off. Appealing to this distinction does not

²⁰Ibid.

²¹Ibid.

²²Temkin [6].

²³O'Neill [4].

help telic egalitarians provide a convincing response to the LDO, since it is implausible to suppose that a state of affairs in which everyone is equally destitute is better in any respect to a state of affairs in which some are better/worse off. However, this distinction can be used to defend the two aforementioned variants of egalitarianism. For it is not implausible to suppose that when there is at least two possible states of affairs, one where a single individual benefits from equality and one where all except this individual benefit from inequality, then the former is better in one respect than the latter. This point can be elaborated further using G.E. Moore's conception of the organic whole. In Moore's account, a complex whole made up of a variety of different elements may possess intrinsic value.²⁴ The value of such a whole bears no relationship to the value of its constituent elements considered in isolation. If we employ this notion, we may say that the organic whole formed by equality and a deriving benefit is intrinsically valuable, whereas the organic whole formed by equality with no deriving benefit lacks any intrinsic value. In the circumstances where equality would marginally benefit just one individual, but inequality would make everyone but him massively better off, the value of the organic whole involving inequality and the benefit many derive from it is greater than the value of the organic whole involving equality and the benefit one person derives from it.²⁵ Crucially, as I have mentioned earlier, there is nothing incoherent in the idea that intrinsic value is dependent upon context. This conception allows intrinsic value to vary from one context to another, provided that in each context its value is grounded in intrinsic properties.

In this essay, I have attempted to attack the LDO from two broad angles. Firstly, in following the method of both Parfit and Temkin, I have identified the LDO as an objection that is grounded in a person-affecting claim. Such a claim, despite its initial rhetorical appeal, can be exposed as inadequate if we widen our considerations to the non-identity problem, proportional justice and the problem of suffering. The latter part of the essay moved away from Parfit's distinction between deontic and telic egalitarianism, and suggested the possibility of two other variants that would also evade the LDO. Through a reassessment of the terms 'non-instrumental', 'intrinsic' and 'extrinsic', we can observe that there is in fact much more room for manoeuvre within the framework of egalitarian thought than Parfit allows. This allows for a more pluralist stance, which in turn accommodates the notion of equality in a variety of different situations, rather than a standpoint based upon an unnecessary di-

²⁴Moore [3].

²⁵Mason [2].

chotomy between instrumental/non-instrumental, and intrinsic/extrinsic values. Thus, the LDO fails to effectively undermine the plausibility of egalitarian claims; however, it teaches us a valuable lesson in that we should be more willing to adopt a pluralist stance concerning equality.

References

- [1] Holtug, N. (1998) 'Egalitarianism and the Levelling Down Objection', in *Analysis*, 58(2): 166–174.
- [2] Mason, A. (2001) 'Egalitarianism and the Levelling Down Objection', in *Analysis*, 61(3): 246–254.
- [3] Moore, G.E. (1903) *Principia Ethica*. Cambridge: Cambridge University Press.
- [4] O'Neill, M. (2008) 'What Should Egalitarians Believe?', in *Philosophy & Public Affairs*, 36(2): 119–156.
- [5] Parfit, D. (2000) 'Equality of Priority?' University of Kansas, Lindley Lecture. Reprinted in (*The Ideal of Equality*). M. Clayton & A. Williams (ed.). Basingstoke: Macmillan.
- [6] Temkin, L.S. (1993) *Inequality*. Oxford: Oxford University Press.

Killing the cat with too much cream? A review of Donaldson and Kymlicka's *Zoopolis* (OUP, 2013, 329 pp.)

Collis Tahzib
Oxford University

Introduction

The animal advocacy movement, Donaldson and Kymlicka claim in their recent book *Zoopolis*, is at an “impasse”.¹ Every year, more and more animals are bred, confined and killed at the altar of satisfying human desires. And we can expect this trend to continue. Animals live, in Patterson’s provocative phrase, in “an Eternal Treblinka”² Why? Because, Donaldson and Kymlicka claim, we treat the animal rights issue as a *moral* issue rather than a *political* one. *Zoopolis* seeks to change this by offering a “political theory of animal rights”. It moves animal rights theory (ART) beyond the familiar terrain of welfare, harm and non-interference and towards exciting political notions like democratic citizenship, sovereignty, migration, colonisation and political participation.

Zoopolis tells the following general story. All animals possess inviolable rights not to be harmed for our benefit (this is traditional ART, and is summarised in Chapter 2). But, *in addition to these basic inviolable rights*, humans *also* bear special relational obligations to different kinds of animals, depending on the nature of the human-animal interaction (this is Donald and Kymlicka’s extra). The authors distinguish between three different kinds of animal and assign each a different kind of political status: (1) “domesticated” animals, who live fully alongside humans, should be considered as *co-citizens* (e.g. dogs, cats, farmed hens); (2) “liminal” animals, who live partially but not wholly with humans, should be considered as *denizens* (e.g. city pigeons, the rabbits in my garden); (3) “wild” animals, who have virtually no contact with humans, should be considered as *sovereign political communities* with rights over themselves and their territory (e.g. sharks, snowy owls). This would have implications at once radical and far-reaching. For instance, granting citizenship rights

¹Donaldson [6], p. 1.

²Ibid. [6] p. 2.

to domesticated animals like household dogs would entail, among other things, the right of dogs to medical treatment, to political representation (if not some form of voting rights in political elections) and to free movement (hence the abolition of leash laws, muzzles and cages except in cases where the animal in question poses a definite and serious threat to themselves or others). Counter-intuitive? Maybe. But we should take our intuitions here with a pinch of salt: after all, at the time, we thought it counterintuitive to extend the franchise to women, to black people, and to other historically subordinated minorities.

In this review I will provide a chapter-by-chapter overview of the authors' arguments before presenting some of my own concerns with their thesis.

Summary of Chapter 2

Chapter 2 deftly summarises and defends ART to date. Here the story is familiar from the work of Francione [7] (2000), Cavalieri [4] (2001), Regan [11] (2003) and others. In essence, traditional ART argues that *all conscious beings, human or animal, are bearers of inviolable rights*. What do we mean by “conscious”? We mean to be sentient, to be aware of pain and pleasure; we mean for there to be an “I” who has subjective experiences of the world, for whom these experiences *matter*, for whom life can go well or badly. We mean, in Smuts's phrase, that there is “someone home”.³ And what do we mean by “inviolable rights”? We mean that a being's basic interests cannot be sacrificed for the good of others. In the case of humans this inviolability is uncontroversial: few believe that it is permissible to kill a human to harvest her organs in order to save five other humans. Yet, in the case of animals, many believe that animal rights may be violated for the good of humans (hence current practices of zoo-keeping, animal experimentation, farming, hunting and so on). Nozick famously described this as “utilitarianism for animals, Kantianism for people”.⁴ So, according to ART, animal rights should be inviolable to the same extent that human rights are.

Critics of ART object to this in various ways, the authors remind us. Some deny that animals really do have consciousness; they deny, that is, that there is “someone home”. However, while it is true that science is still learning about how to cross the phenomenal gap between other minds in order to identify

³Donaldson [6] pp. 24-5.

⁴Donaldson [6] p. 20.

animal consciousness, there are still clear cases either side of a fuzzy line. Thus cockroaches, for instance, probably aren't conscious, while great apes almost certainly are. And, the authors crucially point out, the objection fails to land because the kinds of animals whose rights are most routinely violated are *precisely those kinds of animals whose consciousness is least in doubt*: we domesticate dogs precisely because they can interact with us; we experiment on monkeys precisely because they react similarly to us.⁵

Other critics of ART concede that animals have consciousness but deny that this is sufficient to be a bearer of inviolable rights. Inviolable rights, such critics say, attach to *persons*, where personhood is consciousness plus something extra – where that “extra” may be language, or rationality, or capacity for culture, or long-term planning, etc (p. 26). Thus, objectors say, animals may be conscious but they are not persons, and so are not owed inviolable rights. However, the problem with this objection is that, by making the criteria of personhood more stringent than mere consciousness, many humans (e.g. the comatose, infants, the highly senile, those with congenital mental disabilities) would fail the test of personhood and hence be deprived of rights, while recent scientific studies suggest that great apes, dolphins,⁶ elephants⁷ and wales⁸ would possess the necessary cognitive capacities to pass the test of personhood.⁹ In other words, any attempt by objectors to draw a distinction between conscious beings and persons will cut across the species line, registering some animals as persons and some humans as non-persons. Which is unacceptable – the comatose, the senile, etc., are precisely those most in need of the security of inviolable rights. Hence, to avoid this, rights should be granted according to consciousness and not some more stringent notion of personhood.

In response, the critic of ART may try to make personhood coextensive with humanity by saying that personhood requires consciousness plus the “species potential” for language, rationality, etc. But the problem with this is that, by tweaking the boundaries of personhood such that they neatly coincide with being human, the account collapses into “speciesism” – discrimination based on morally arbitrary biological facts about being a member of *homo sapiens*, which seems to be a form of discrimination akin to racism or sexism.

⁵Ibid. [6] p. 31.

⁶White [12] 2007.

⁷Poole, [10] 1998.

⁸Cavalieri, [3] 2006.

⁹Donaldson [6] p. 26.

Summary of Chapter 3

Chapter 3 contains all the new theory. It extends and supplements contemporary ART by drawing on citizenship theory. Citizenship theory says that rights should be differentially allocated according to citizenship status. So, for instance, citizens may have special rights that are denied to foreigners (e.g. the right to stay permanently in the country, the right to participate in political elections, and so on). And these citizenship rights, Donaldson and Kymlicka claim, apply in the case of animals too – so domesticated animals, for instance, ought to be seen as co-citizens.

Of course, there is an immediate objection to this: surely animals aren't citizens because citizenship implies active agency and participation (voting, engaging in political debate, etc) that animals are in principle incapable of? The authors respond to this objection in two ways. First, animals *are* capable of certain kinds of political agency. To be sure, they cannot exhibit their agency in the way humans do (by entering the polling booth, by public debate, etc). But, drawing on insights from contemporary disability theory, the authors argue that there are *other ways* of expressing political agency. One is the notion of guardianship: “guardians” or “trustees” could vote *in the interests of animals*. Another is to use new models of “supported decision-making” or “dependent agency” for what Wong calls “non-communicating citizens” – which involves eliciting a person's sense of their good, their interests and their preferences through nonverbal communication such as pointing, making sounds, laughing, and choosing to reside in one place rather than another (“voting with one's feet”).¹⁰ Such supported decision-making is hardly controversial or new: after all, we often use assisted agency for children, for the disabled, for those temporarily incompetent due to illness, for immigrants who need translation, and so on.¹¹ So there seems to be no principled reason why such supported decision-making should not be provided to animals too. And second, *even if animals are not capable of exhibiting democratic political agency*, it does not follow that they do not possess citizenship rights. For if we make the ability to exercise political agency a necessary condition of citizenship rights then we are led, perversely, to exclude from citizenship children and the mentally disabled (who are equally unable to participate in political debate).

¹⁰Ibid. [6] pp. 59, 66.

¹¹Ibid. [6] p. 60.

Summary of Chapters 4-7

Chapters 4-7 apply the theory developed in Chapter 3 to the cases of domesticated animals, wild animals and liminal animals. Chapters 4 and 5 treat the case of domesticated animals. Domesticated animals – by which is meant animals such as dogs, sheep and chickens who have for various reasons been selectively bred by humans – should, the authors argue in these chapters, be recognised as *citizens* and accorded all the attendant citizenship rights. How could animals be citizens? To answer this, we need first to clarify what we mean by citizenship. *Citizenship* implies being an active participant in and co-author of a community’s laws, institutions and culture. To do this, it is often said (following Rawls) that three basic capacities are required: (1) the capacity to have a subjective good and to communicate that good; (2) the capacity to cooperate and to comply with social norms; and (3) the capacity to participate in the making of laws.¹² Now, these three basic capacities *can* be interpreted in what the authors call a “rationalist” or “intellectualist” way: to have *reflectively* endorsed one subjective good rather than another; to *understand the reasons* for compliance with social norms; to participate in law-making through *rational debate, advocacy* and so on.¹³ But, as recent disability theorists have argued, by interpreting the basic capacities of citizenship in these highly rationalist ways, we exclude from citizenship children, the severely mentally disabled, those with dementia, and so on. To avoid this abhorrent conclusion, we must, disability theorists argue, reconceive these three basic capacities in more inclusive, less intellectualist ways. At the heart of these more inclusive reconceptualisations is the notion of “dependent agency”, which is exercised through loving relationships with particular others in whom they trust and who have the skills to look for, recognise, and nurture the expression of their agency (p. 104). For instance, Kittay [9] (2001) illustrates how a severely disabled person communicates to her caregiver that first basic capacity, her subjective good, through “a glint in the eye at a strain of familiar music”, “a slight upturn of the lip ... when a favourite caregiver comes along”, “a look of joy in response to the scent of perfume”.¹⁴

The authors then apply these new, more inclusive, less rationalist conceptions of citizenship and agency to domesticated animals. If the severely disabled meet the basic requirements of citizenship despite not being capable of rational

¹²Ibid. [6] p. 103.

¹³Ibid. [6] pp. 103-4.

¹⁴Ibid. [9] p. 264.

communication, then why not domesticated animals too? Let us consider each of the three basic capacities required for citizenship.

(1) *Capacity to have and to communicate a subjective good?* Animals have a whole repertoire of gestures, movements and signals to express their good (in the sense of their interests, preferences and desires): meowing in front of the fridge to ask for food, bowing to invite you to play, bringing over the leash to indicate walk time, whimpering to express distress or pain, etc.¹⁵ Once we start “expecting agency, looking for agency, and enabling agency”¹⁶ we see it all around: domesticated animals have a good, and can *communicate* that good, even though they cannot *rationally reflect* on that good.

(2) *Capacity to cooperate and comply with social norms?* Research in the science of animal behaviour by Bekoff and Pierce [1] (2009) and de Waal [5] (2009) is increasingly revealing animals’ capacity for reciprocity, justice, altruism and empathy. For instance, in a famous study by Brosnan and de Waal [2] (2003), two capuchin monkeys are each perfectly happy to perform a certain task when each is rewarded with cucumber; but, when one is rewarded with cucumber and the other with grapes (which capuchins prefer), the cucumber-receiving capuchin becomes furious and throws the cucumber back at the researcher in protest. Capuchin monkeys, then, have a sense of *fairness*. In another example illustrating capacity to participate in cooperative society, Mason describes how Lulu, a pig, saved her human companion, Joanne Altsmann: sensing that Joanne was unwell, Lulu “forced her body through a dog door, scraping herself and drawing blood in the process... ran out to the road and lay across it until a car stopped, then led the driver back to the kitchen where Altsmann had suffered a heart attack” (p. 119).

(3) *Capacity for political participation?* One form of political participation is *sheer presence*. Thus dogs – especially guide dogs, who assist people with disabilities – in public in Europe are their own advocates against, say, North American tourists whose laws back home prohibit such animal presence in public spaces on grounds of public health.¹⁷ Another form of political participation is protest, dissent and revolt. A fascinating study in this respect comes from Hribal, who in his book *The Hidden History of Animal Resistance* explores the political agency of working animals. Hribal argues that “part of the explanation for the rapid transition from horse power to the internal com-

¹⁵Donaldson [6] pp. 108-9.

¹⁶Ibid. [6] p. 109.

¹⁷Donaldson. [6] pp. 113-4.

bustion engine in the early twentieth century lies in industrial management's desire to be rid of a disruptive [animal] workforce who regularly challenged their working conditions" through escape attempts, violence, strikes, destruction of machinery and general obstreperousness.¹⁸ When humans do these things, we call it political protest; why, then, are we so reluctant to attribute political agency when animals do so too? The same protest, Hribal argues, also occurs among recalcitrant zoo and circus animals. Thus, once we interpret the three basic capacities in an expansive, less rationalist way – a way that recognises the disabled as citizens too – we find much evidence that domesticated animals possess the requisite capacities for citizenship.

Chapters 6 and 7 apply the same kind of argument to the case of wild animals and liminal animals; so to avoid repetition we will treat these cases in less detail. Chapter 6 argues that wild animals – defined as those animals that maintain a self-sufficient existence apart from human contact and human settlements (e.g. deep sea anglerfish or Saharan desert larks) – ought to be thought of as forming "sovereign communities whose relations with human communities should be regulated by norms of international justice".¹⁹ Hence the authors argue for checks on direct violence (like hunting, fishing, killing animals as part of wildlife management schemes), on habitat loss through human encroachment and urbanisation (which they interestingly compare to colonisation), and on spillover harms from shipping lanes, roadways, pollution, skyscrapers that disrupt the migratory flight paths of birds, and so on.²⁰ Chapter 7 argues that "liminal animals" – defined as non-domesticated animals (like deer, squirrels, rats) that live in our midst without being under direct human care – should have denizenship status. Denizenship is a weaker relationship than citizenship and carries a reduced set of rights and responsibilities – examples of denizens in human political communities include migrants, refugees and isolationist communities such as the Amish.

Five Brief Objections

Before closing, let me present five brief objections of my own to the arguments in *Zoopolis*. First: the authors' three categories – citizenship, denizenship and sovereign subject – lean on a statist conception of international justice. On the

¹⁸Hribal [8] p. 115.

¹⁹Donaldson [6] p. 157.

²⁰Ibid. [6] pp. 156-7.

statist conception, we owe different rights to a co-citizen than we do to a non-citizen denizen than we do to a sovereign foreign political community. And these group-differentiated rights roughly map across in the case of the rights owed to different kinds of animals. However, cosmopolitanism rejects these group-differentiated rights; on the cosmopolitan view, all persons qua persons (rather than qua citizens or any other political status) are owed equal rights. This creates problems for the authors' thesis, yet they fail to engage with this cosmopolitan alternative.

Second: the authors argue that wild animals should be treated as a kind of sovereign state, with its own right to self-determination and non-interference by outsiders. But this raises a problem of "humanitarian intervention". As the authors concede, intervention in sovereign states may be justified in the human case in instances of natural disaster or internal chaos (e.g. a "failed state" or genocide, where rights are routinely violated).²¹ But are not the majority of wild animal communities "failed states", where lions will pick off gazelles with impunity and with no regard for gazelle rights, and where whole species are systematically devoured by animals higher on the food chain? If this is so, then we would seem, very counterintuitively, to be committed to constant and intensive humanitarian interventions in nature – if you thought Iraq was bad, think again.

Third: given the sheer quantity of animals, and their rapid rates of reproduction, it is possible that as animal interests are increasingly factored into political decision-making, as the authors encourage, we could see a tyranny of the *animal* majority! Of course, this need not defeat vision of *Zoopolis*; but surely it at least requires addressing.

Fourth: like much other writing on animal rights, the authors lean heavily on a move known as "the argument from marginal cases", which runs as follows: if we are to exclude animals from citizenship on the grounds that they lack certain powers requisite for citizenship then we must on pain of contradiction also exclude the comatose, the severely disabled, etc (the so-called "marginal cases") who also lack such powers; but this is perverse; so to avoid excluding these marginal cases we must modify (in particular, relax) the powers requisite for citizenship in such a way that the marginal cases are included; but these modifications (i.e. relaxations) will end up including animals too. However, the argument from marginal cases (AMC) is not unproblematic and the authors' thesis will inherit any objections to AMC (e.g. that of trying to jus-

²¹Ibid. [6] p. 182.

tify speciesism, that of biting the bullet and denying citizenship to the severely disabled, or that of arguing that AMC can lead to the denigration and maltreatment of marginal human beings), objections that the authors fail to mention let alone defuse.

And my fifth objection is strategic rather than philosophical: arguing that animals are in some cases citizens may in fact end up being counterproductive, for it may alienate the unconverted, who having not yet accepted even the most minimal claims of traditional animal rights theory are hardly likely to show enthusiasm for Kymlicka and Donaldson's new, extended theory. Nor, I think, are the authors' claims likely to capture the popular imagination or enter mainstream political rhetoric. So they may hinder rather than help the animal advocacy movement by *killing the cat with too much cream*, if you'll forgive the pun. Animal advocates, then, might have more success, politically and strategically speaking, by mobilising behind the most urgent, vivid and egregious animal abuses (the carnage of animal slaughterhouses that support our demand for meat consumption, mass animal exterminations as a form of animal population control, etc.) before trying to win support for animals' positive rights to mobility or political representation.

In conclusion, *Zoopolis* is a strikingly original and carefully argued addition to the animal advocacy literature. Donaldson and Kymlicka are right to note that ART to date has focused overly on negative rights for animals rather than the inevitable question of the kind of relations and interactions that we *should* in fact have with animals. In this sense, the "political turn" is welcome. But, strategically, *Zoopolis*'s political turn may do more harm than good: representing pets as fellow citizens may make the animal advocacy movement appear fanciful; it alienates the unconverted; it kills the cat with too much cream.

References

Bibliography

- [1] Bekoff, M. & Pierce, J. (2009) *Wild Justice; The Moral Lives of Animals*. Chicago: University of Chicago Press.
- [2] Brosnan, S. & de Waal, F. (2003) "Monkeys reject unequal pay" in *Nature* 425:297-299.

- [3] Cavalieri, P. (2006) "Whales as persons" in Kaiser, M. & Lien, M. (eds) *Ethics and the politics of food*. Wageningen: Wageningen Academic Publishers.
- [4] Cavalieri, P. (2001) *The Animal Question: Why Nonhuman Animals Deserve Human Rights*. Oxford: OUP.
- [5] de Waal, F. (2009) *The Age of Empathy: Nature's Lessons for a Kinder Society*. (Toronto: McClelland & Stewart).
- [6] Donaldson, S. & Kymlicka, W. (2013) *Zoopolis: A Political Theory of Animal Rights*. Oxford: OUP.
- [7] Francione, G. (2000) *Introduction to Animal Rights: Your Child or the Dog?* Philadelphia: Temple University Press.
- [8] Hribal, J. (2006) *Fear of the Animal Planet: The Hidden History of Animal Resistance*. Oakland: Counter Punch Press and AK Press.
- [9] Kittay, E. (2001) "When Caring is Just and Justice is Caring: Justice and Mental Retardation" in *Public Culture* 13:557-79.
- [10] Poole, J. (1998) "An Exploration of a Commonality between Ourselves and Elephants" in *Etica & Animali* 9:85-110.
- [11] Reagan, T. (2003) *Animal Rights, Human Wrongs: An Introduction to Moral Philosophy*. Lanham, MD: Rowman & Littlefield.
- [12] White, T. (2007) *In Defense of Dolphins: The New Moral Frontier*. (Oxford: Blackwell)

The Dialectic of the Formal and the Intuitive: Kurt Gödel's Proof of the Ideality of Time

Lukas Clark-Memler
University of Edinburgh

Those affections which we represent to ourselves as changes, in beings with other forms of cognition, would give rise to a perception in which the idea of time, and therefore also of change, would not occur at all. — Immanuel Kant¹

Relativity theory has furnished a very striking illustration, in some sense even a verification, of Kantian doctrines. — Kurt Gödel²

Kurt Gödel's 1949 paper, 'A Remark About the Relationship Between Relativity Theory and Idealistic Philosophy', is the logician and mathematician's sole published foray into the philosophy of time. Gödel's 'remark' is that our intuitive conception of time, as something that objectively lapses,³ is not true of reality itself; relativity theory proves there is no physical correlate for our experience of time. Derived from his solutions to Einstein's field equations for the General Theory of Relativity (GTR), Gödel constructs a cosmological model to argue against the reality of time, aligning himself firmly in the tradition of temporal idealism.⁴ Gödel's thesis may appear shocking, especially considering the brevity of his paper,⁵ but the underlying physical proof, and force of his modal argument, makes it difficult to deny.

Hao Wang, Gödel's principal interlocutor, writes that Gödel's philosophical work was largely motivated by "the dialectic of the formal and the intuitive".⁶

¹ *Critique of Pure Reason* p. 54.

² 'Some Observations About the Relationship Between Theory of Relativity and Kantian Philosophy,' p. 247. Henceforth referenced as 'Observations.'

³ Gödel [9] pp. 557–558.

⁴ *Ibid.* Gödel cites Parmenides and Kant as the principal influences of his idealism; we can add to this Plato, Leibniz, McTaggart, and Husserl.

⁵ Gödel's 'Remark' is only six pages long, and his 'Cosmological Solutions' to GTR is a mere four pages long. If "brevity is the soul of wit" (*Hamlet*, Act II, Scene ii) then Gödel is certainly a wit *par excellence*.

⁶ Wang [24] p. 8. Because of the lack of Gödel's philosophical publications, Wang's accounts of Gödel's musings on time are of immense value.

We witness this ‘dialectic’ in Gödel’s mathematical work also – his Incompleteness Theorems show the tension between formalism and intuition. For want of a better term, I will refer to this dialectic as the ‘Gödel Project’. We can view Gödel’s cosmological equations as a systematic attempt to represent an intuitive concept with a formal model. In this light, we see that Gödel’s argument for the ideality of time (the unreality of time in the physical universe) is based on the inability of relativity theory to formalise intuitive time.

This inquiry will subject to analysis Gödel’s argument and the consequence of his proofs. By considering Gödel’s influences and philosophical views (in particular, his mathematical realism and Platonism); the ontology of time (i.e. McTaggart’s A-series and B-series) and the intuitive conception of time; Gödelian universes and his modal argument; we come to the conclusion that time as we experience it – objectively lapsing or flowing – has no place in a relativistic universe. Thus, Gödel successfully proves the ideality of time, and the physical inapplicability of intuitive time. Finally, I will discuss how the Gödel Project highlights inconsistencies between the results of Gödel’s GTR solutions and Incompleteness Theorems.

In his autobiography, Bertrand Russell describes Gödel as an “unadulterated Platonist”.⁷ While Russell certainly meant this disparagingly (as Gödel had used his ‘Platonism’ to critique and disprove Russell’s own *Principia Mathematica*) Gödel did consider himself a ‘mathematical realist’ believing that, “the objects and theorems of mathematics are as objective and independent of our free choice and our creative acts as is the physical world”.⁸ Gödel’s brand of mathematical Platonism is distinctly Leibnizian; like Leibniz, Gödel’s philosophical arguments are based on mathematics.⁹ Later in his life, Gödel became interested in Edmund Husserl’s phenomenology. Phenomenological analysis begins with “the complete exclusion of every assumption, stipulation, or conviction concerning objective time”.¹⁰ Following Husserl, and to a lesser extent Kant, Gödel espoused a strong temporal idealism, much to the chagrin of his

⁷Russell [21] p. 356.

⁸Gödel [6] p. 312 fn. 17.

⁹Van Atten & Kennedy [23] p. 275–281. We see Leibniz’s influence most clearly when Gödel discusses his ‘conceptual realism’: he writes that “Concepts have an objective existence” (Wang, *Logical Journey* p. 318).

¹⁰Husserl [10] p. 4.

analytical and positivist peers.¹¹

In his ‘Remark’, Gödel characterises time as “that mysterious and seemingly self-contradictory being which, on the other hand, seems to form the basis of the world’s and our own existence”.¹² By ‘self-contradictory’, Gödel is referring to John McTaggart’s seminal analysis of time, in which the very concept of time is proven to be logically incoherent.¹³ While Gödel’s project is physical rather than logical (this sets him apart from the idealist tradition of a priori arguments for the unreality of time), we should consider McTaggart’s analysis in order to clarify the meaning of time.

There are two dominant, and competing, theories of time: A-theory (tensed) and B-theory (tenseless). According to A-theory, time is ordered by determinations of past, present and future; neither the past nor future exist, the present ‘now’ has ontological privilege.¹⁴ The three monadic predicates of the A-series capture the concept of change in the infinite succession of layers of existence coming into being (temporal ‘becoming’). The B-series is instead ordered by fixed temporal relations of ‘earlier’ and ‘later’, denying objective existence to the present; all terms in the eternal temporal series are ontologically on par. Modern physics has appropriated McTaggart’s B-theory with the ‘block universe’¹⁵ model: the two dyadic predicates exist tenselessly in a four-dimensional space-time manifold.¹⁶ B-theorists treat time like “a Platonic object whose members do not come into existence successively”¹⁷ where the present ‘now’ is relativised depending on the subject; any talk of change or passage is unreal (i.e. not a property of the physical universe).

¹¹In fact, Gödel’s idealism (the *bête noire* of the analytical tradition) may have been what stopped him from publishing more extensively on the topic of time.

¹²Gödel [9] p. 557. We are reminded of Augustine’s distress when attempting to describe time in his *Confessions*: “For what is time? Who can easily and briefly explain it? ... If no one ask of me, I know; if I wish to explain to him who asks, I know not.” (Book 11, Ch. 14)

¹³McTaggart [15]. Note that McTaggart is making a metaphysical claim, not merely a linguistic one.

¹⁴Olson [17] pp. 440–441.

¹⁵This characterisation is first attributed to William James, from his 1908 Hibbert Lecture.

¹⁶Oaklander [16] p. 256. This is the conception of time first proposed by the Eleatic School.

¹⁷Yourgrau [27] p. 11.

There are problems inherent in both theories,¹⁸ though B-theory's ontological neutrality seems to give it at least a *prima facie* advantage in the face of relativity theory. In fact, Gödel follows McTaggart in assuming that change (or 'lapse') is an essential aspect of time, and simply does not consider B-theory a valid temporal series.¹⁹ Gödel argues that change is at the heart of our experience of time – our use of tensed language supports his claim that our temporal intuition is dynamic²⁰ – and that “change becomes possible only through the lapse of time. The existence of an objective lapse of time... [means] that reality consists of an infinity of layers of 'now' which come into existence successively.”²¹

This certainly sounds like the A-series, however, this theory faces problems when relativity is introduced. The most significant effect on physics of the Special Theory of Relativity (STR) is the denial of absolute simultaneity and a privileged reference point: “The subjectivist consequences of relativity theory... [is that it] does not permit of any objective lapse of time or objective change.”²² Since A-theory requires an ontologically privileged reference point (from which to divide events into the three monadic predicates) it suffers in the face relativity. And since Gödel rejects B-theory, time must be ideal.²³

¹⁸McTaggart [15] pp. 471–473; Olson [17] pp. 446–447. To briefly characterise the self-contradictory nature of the A-series: the past, present and future are mutually contradictory, ‘incompatible determinations’, yet according to A-theory, all events in time must possess the three contradictory properties; since an event that is in the *future* will become *present* and then *past*.

¹⁹Einstein, on the other hand, subscribed to B-theory time for much of his life. Karl Popper visited Einstein at Princeton in 1950, and recalls a conversation he had with him about the nature of time: “I tried to persuade him [Einstein] to give up his determinism, which amounted to the view that the world was a four-dimensional Parmenidean block universe in which change was a human illusion, or very nearly so. (He agreed that this had been his view, and while discussing it I called him ‘Parmenides’.) I argued that if men ... could experience change and genuine succession in time, then this was real.” (*Unended Quest* pp 129–130).

²⁰Even if we disagree with Gödel that our temporal intuition is A-theoretic (and contend that we experience time as a static, relational series), his cosmological model rules out the B-series: closed time-like curves and the possibility of backwards time travel entirely disrupt the eternal dyadic predicates.

²¹Gödel [9] p. 558.

²²Gödel [5] p. 237.

²³Gödel's argument for the ideality of time according to STR: (1) Time requires change, (2) Change requires an objective lapse, only possible with A-theory, (3) STR rejects A-theory, (4) Therefore, time is ideal (Bourne, *A Future for Presentism* p. 206). Note that Bourne finds Gödel's argument unsatisfactory, and continues to defend the A-theory.

If STR was all Einstein had to say about our universe then Gödel's defence of temporal idealism would be complete; however, the development of GTR brought with it a possible vindication for tensed time.

The presence of matter in GTR offers a privileged status to a certain frame of reference: observers that follow the "mean motion of matter" can consider their "time as the true one, which lapses objectively".²⁴ A-theory rears its head once more, and it seems that the conditions of absolute time are met in most cosmological models of GTR. Gödel's solution to GTR, however, involves a non-expanding, rotating universe: "The compass of inertia in these worlds rotates everywhere relative to matter",²⁵ making it impossible to derive a consistent global time. Much has been written about the physical parameters of a Gödelian universe (GU), and no decisive refutation has arisen.²⁶ Assuming that GU is a physically sound model, A-theory is unequivocally defeated.

A possible response is that Gödel did not prove the unreality of time, but only the unreality of an objective lapse of time; time survives, albeit in a relativised state, and certainly not in any tensed series. While this is true, Gödel writes that, "the concept of existence... cannot be relativized without destroying its meaning".²⁷ He considers relativised time "absurd" and wholly different from our intuitive conception of time.²⁸ Here we see the relevance of the Gödel Project, in the tension between the formal and the intuitive. As a 'conceptual realist' Gödel is unwilling to compromise or adapt his notion of intuitive time.²⁹ If GTR cannot formalise full-blooded intuitive time, then time must be ideal.

²⁴Gödel [9] p. 559, emphasis added. In his 'Solutions' Gödel writes, "All cosmological solutions with non-vanishing density of matter known at present have the common property that... they contain an 'absolute' time coordinate." (447)

²⁵Gödel [5] p. 248 fn. 7.

²⁶Refuting GU would require proving that Gödel's solution to GTR is either A) incorrect, logically or mathematically; or B) has no bearing on our physical universe.

²⁷Gödel [9] 558 fn. 8.

²⁸Ibid.

²⁹The key question here is: does Gödel's conceptual realism include the concept of 'time'? It does seem that his conceptual realism extends beyond mathematical objects, as he explicitly writes, "I was a conceptual and mathematical realist since about 1925." (Hao Wang, *Reflections on Gödel* p. 20.) While he separates the two categories, he is ambiguous in regards to the extent of the purview of his conceptual realism. We should note that temporal idealism and conceptual realism complement each other well.

Even though Gödel did not consider tenseless time valid (since it does not correspond to intuitive time), it is interesting to see how B-theory fares in the face of GU. While the absence of an objective lapse and frame of reference in GU does not damage the B-series, the existence of closed time-like curves (CTCs) does. CTCs are a remarkable feature of GU that seem to allow (at least conceptually) for backwards time travel – by offering the traveller a curved, continuous worldline that, when followed, eventually leads back to where it began.³⁰ The possibility of CTCs disrupts the fixed temporal order (of ‘earlier-later’ relations) necessary to B-theory: “If P, Q are any two points on a world line of matter, and P precedes Q on this line, there exists a time-like line connecting P and Q on which Q precedes P.”³¹ And so B-theory collapses into a jumbled heap of temporal relations.

The physical peculiarity of GU leads detractors to claim that Gödel’s proof of the ideality of time has no bearing on *our* universe; ‘So much for time in GU’, the temporal realist responds. While we have substantial astrophysical evidence that our universe is *not* a rotating static block (for example, the observed redshift of distant galaxies), we should remember that Gödel’s solution to GTR complies with the laws of nature and cosmological parameters of our universe; in other words, GU is a nomological (and not a mere logical) possibility. The fact that GU is counter-intuitive does not discount its nomological compatibility, which only requires that it comply with a particular physical theory. Nomological assessment is relative to a certain theoretical background; we should ask, ‘If we accept GTR as a physical theory of our universe, what other cosmological models are nomologically accessible?’

The essence of the modal argument is that time must objectively lapse in all physically possible worlds for it to exist in any of them; we must be able to identify absolute becoming in all nomologically compatible universes. In Gödel’s words: “The possibility of a determination of an absolute world time, where it exists at all, is certainly not due to the laws of nature (which are satisfied in all cosmological solutions), but only to the special distribution and motion which matter has in those instances.”³² Gödel argues that the differences between GU and our universe are only contingent – the ‘distribution and motion’ of matter – and that time, if it were to exist in an absolute way,

³⁰“A trip along a CTC is a journey into the past.” (Barry Dainton, *Time and Space* p. 316.)

³¹Gödel [8] p. 447.

³²Gödel [5] p. 238.

cannot depend on mere contingency.³³ While we may disagree about fundamental and contingent facts, GU remains nomologically possible, and thus should be taken seriously. Gödel’s modal argument can only be refuted by denying the given physical theory (i.e. GTR).

Still, we do *experience* time in our universe, so if Gödelians were phenomenologically distinct from us, that could explain the temporal differences. Yet the “lawlike compossibility” of the two universes tends on the side of phenomenological parity.³⁴ Gödel notes that, “In whatever way one may assume time to be lapsing, there will always exist possible observers to whose experienced lapse of time no objective lapse corresponds... But, if the experience of the lapse of time can exist without an objective lapse of time, no reason can be given why an objective lapse of time should be assumed at all.”³⁵ Time in GU is ideal, and since GU is nomologically possible, and Gödelians experience time as we do, we have no sufficient reason to think that time is real in our universe, and are forced to accept the ideality of time. Such is the power of the modal argument.

It is clear that neither the A-series nor the B-series can exist in a relativistic universe, yet Gödel’s GTR equations still use the time variable ‘*t*’; what, then, remains of time? Relativity theory interprets time spatially and geometrically: *t* is simply the temporal component of four-dimensional space-time. STR and GTR prove that time and space do not have independent existence. We find that there is ontological incompatibility between time (*t*) and Time (capital ‘T’: the objective, independent concept of our intuition). It is ‘Time’ that Gödel proves is ideal; space-time is *the* only temporal feature of our universe. While we subjectively experience ‘Time’ – lapsing and changing – it has no physical correlate.

The dialectic of the formal and the intuitive motivates both Gödel’s cosmological solutions to GTR and his famous Incompleteness Theorems. But, curiously, the Gödel Project yields inconsistent results. Gödel’s Incompleteness Theorems prove that purely formal systems (i.e. Hilbert’s program) are unable to capture all mathematical truths; any formal proof is essentially incomplete. The inadequacy of formal methods in capturing all intuitive arithmetic con-

³³Contingent facts about our universe generally involve the distribution of matter, for example, how many planets there happen to be in our solar system.

³⁴Savitt [22] pp. 466–467. Savitt writes, “The Gödelians will... have experience of an objectively lapsing time, just as we do... Why should our experience be taken as any indication that in our world time does objectively lapse?”

³⁵Gödel [9] p. 561.

cepts leads to the refutation of a complete formalism. Yet in regards to time, the inability of GTR to incorporate intuitive time demonstrates the unreality of time; it is not Einstein's theory that is incomplete, but rather an error of intuition. We are left wondering why Gödel's cosmological model did not lead him to the conclusion that relativity theory was flawed and limited, instead of considering his results a "striking illustration, in some sense even a verification, of Kantian doctrines".³⁶ In the former case, the tension between the formal and the intuitive reveals the flaws in formalism, while the latter case highlights problems with our intuition. It is beyond the scope of the present inquiry to explain the asymmetry between Gödel's interpretation of the results of GTR and the Incompleteness Theorems, but this is an inconsistency that certainly deserves further consideration.³⁷

Gödel's 'Remark' grapples with the ontology of time, a problem that haunts the canon of Western thought – from Parmenides, to Leibniz, to Kant. He shows us that neither A-theory nor B-theory time can be reconciled with GTR; it would take a revision of Einstein's theory – perhaps in the field of quantum gravity, leading to a new 'Theory of Everything' – in order to allow for time to exist objectively in our universe. By first distinguishing the formal concept of time from the intuitive, then constructing a cosmological model in which intuitive time cannot be captured, Gödel demonstrates the fundamental incompatibility of intuitive time with a relativistic world. His modal argument forces us to take his GTR solutions seriously, and we are led to the conclusion that our subjective experience of time does not correspond to any physical reality. Time is ideal. The last word belongs to Gödel: "What remains of time in relativity theory... neither has the structure of a linear ordering nor the character of flowing or allowing of change. Something of this kind, however, can hardly be called time. So Kant's view about time *has* been confirmed by relativity theory."³⁸

³⁶Gödel [5] p. 247. See the epigraph of this essay for a description of Kant's temporal idealism.

³⁷While this is more of a curious inconsistency than philosophical problem, it is surprising how little attention this has been given. The lack of Gödel's published work in this area presents a difficulty in solving the inconsistency of his thought. Palle Yourgrau discusses the asymmetry at length in his work on Gödel.

³⁸Gödel [5] p. 236, emphasis in original.

References

- [1] Augustine. (2002) *Confessions*, A. C. Outler (trans.). Mineola, NY: Dover Publications.
- [2] Bourne, C. (2006) *A Future for Presentism*, New York: Oxford University Press.
- [3] Broad, C.D. (1955) 'Time', in J. Hastings (ed.), *Encyclopaedia of Religion and Ethics Vol. 12*. New York: Scribner.
- [4] Dainton, B. (2001) *Time and Space*. Durham: Acumen.
- [5] Gödel, K. (1995) 'Some Observations About the Relationship Between Theory of Relativity and Kantian Philosophy', in S. Feferman, W. Goldbarb, C. Parsons & R.M. Solovay (ed.), *Collected Works Vol. III: Unpublished Essays and Lectures*. New York: Oxford University Press.
- [6] Gödel, K. (1995) *Collected Works Vol. III: Unpublished Essays and Lectures*. S. Feferman et al.(ed.). New York: Oxford University Press.
- [7] Gödel, K. (1990) *Collected Works Vol. II: Publications 1938–1974*, S. Feferman, J.W. Dawson Jr., S.C. Kleene, G.H. Moore, R.M. Solovay & J. van Heijenoort (ed.). New York: Oxford University Press
- [8] Gödel, K. (1949) 'An Example of a New Type of Cosmological Solutions of Einstein's Field Equations of Gravitation', in *Reviews of Modern Physics* 21: 447–450.
- [9] Gödel, K. (1949) 'A Remark About the Relationship Between Relativity Theory and Idealistic Philosophy', in P. A. Schilpp (ed.), *Albert Einstein: Philosopher-Scientist*. La Salle, IL: Open Court.
- [10] Husserl, E. (1964) *The Phenomenology of Internal Time-Consciousness*, M. Heidegger (ed.), J.S. Churchill (trans.). Bloomington, IN: Indiana University Press.
- [11] James, W. (1909) *A Pluralistic Universe: Hibbert Lectures* London: Longman.
- [12] Kant, I. (2007) *Critique of Pure Reason*, N.K. Smith (trans.). London: Palgrave Macmillan.
- [13] Malament, D. (1984) 'Time Travel in the Gödel Universe', in *Philosophy of Science Association, Volume II: Symposia and Invited Papers*: 91–100.

- [14] Martin, D. (2010) 'Gödel's Conceptual Realism', in S. Feferman, C. Parsons & S.G. Simpsons (ed.), *Kurt Gödel: Essays for His Centennial*. Cambridge: Cambridge University Press.
- [15] McTaggart, J. (1908) 'The Unreality of Time', in *Mind* 68: 457–474.
- [16] Oaklander, L.N. (1992) 'Thank Goodness It's Over', in *Philosophy* 260: 256–258.
- [17] Olson, E. (2009) 'The Passage of Time', in R. Le Poidevin, P. Simons, A. McGonigal & R. Cameron (ed.), *The Routledge Companion to Metaphysics*. London: Routledge.
- [18] Parsons, C. (1995) 'Platonism and Mathematical Intuition in Kurt Gödel's Thought', in *The Bulletin of Symbolic Logic* 1: 44–74.
- [19] Popper, K. (1976) *Unended Quest: An Intellectual Autobiography*. La Salle, IL: Open Court.
- [20] Potter, M. (2001) 'Was Gödel a Gödelian Platonist?', in *Philosophia Mathematica* 3: 331–346.
- [21] Russell, B. (1968) *The Autobiography of Bertrand Russell*. London: Allen and Unwin.
- [22] Savitt, S. (1994) 'The Replacement of Time', in *Australasian Journal of Philosophy* 72: 463–474.
- [23] Van Atten, M. & Kennedy, J. (2010) 'On the Philosophical Development of Kurt Gödel', in S. Feferman, C. Parsons & S.G. Simpsons (ed.), *Kurt Gödel: Essays for His Centennial*. Cambridge: Cambridge University Press.
- [24] Wang, H. (1996) *A Logical Journey: From Gödel to Philosophy*. Cambridge, MA: MIT Press.
- [25] Wang, H. (1995) 'Time in Philosophy and in Physics: From Kant and Einstein to Gödel', in *Synthese* 102: 215–234.
- [26] Wang, H. (1987) *Reflections on Kurt Gödel*. Cambridge, MA: MIT Press.
- [27] Yourgrau, P. (1991) *The Disappearance of Time: Kurt Gödel and the Idealistic Tradition in Philosophy*. Cambridge: Cambridge University Press.

Michael Smith's 'Moral Problem' and a Defence of Metaethical Externalism

Laura Nicoara

University of Glasgow

Introduction

There are three doctrines which reflect the ways in which we intuitively think and speak about morality. The objectivity thesis states that moral judgments are beliefs about objective facts. Moral internalism views moral judgments as intrinsically motivating: to judge that Φ -ing is right is to necessarily be motivated to Φ . Finally, the Humean Theory of Motivation claims that moral motivation is generated by an appropriate belief–desire pair, where no belief entails a desire. These three claims, although separately plausible, appear to be jointly inconsistent.¹ If moral judgments are beliefs that motivate intrinsically and necessarily, then moral motivation is possible without the existence of a motivating desire (anti-Humeanism). If moral judgments are beliefs, and beliefs only motivate in the presence of an appropriate desire, then moral judgments cannot be intrinsically motivating, and motivation must depend on something else, such as the nature and dispositions of agents (metaethical externalism). Finally, if moral judgments are intrinsically motivating, and beliefs are not sufficient for motivation, then moral judgments are not beliefs (non-cognitivism).

This paper will argue that the best solution to the moral problem is externalism. It will do so by starting from Michael Smith's argument in his 1995 book *The Moral Problem*, in which he proposes an internalist, rationalist view of moral motivation. In the first section of the article, I show that Smith's objections to externalism are inconclusive. Next, I argue that Smith's own rationalism about moral motivation leads back to an externalist perspective. Finally, I follow Smith's model of conceptual analysis of rightness and attempt provide a more adequate analysis, one which can accommodate externalism about moral motivation.

¹Smith [4] p. 12.

1 The Moral Problem and the externalist solution

The internalist doctrine seems to capture the intuition that we normally speak and think as though having a moral conviction that Φ -ing is right is sufficient for motivating one to Φ .² We would find it odd if someone were to sincerely believe a moral claim such as ‘Stealing is wrong’ and at the same time to engage in acts of stealing while in a fully rational state, without being coerced, etc. The externalist, however, will claim that, while such a conjunction between moral judgment and motivation is often expected of, and present in, moral agents, its absence is not a conceptual impossibility in a rational person, hence what has become known as ‘the amoralist challenge’.³

Smith’s first argument against externalism consists in a rejection of the amoralist as a conceptual possibility. On its own this is inconclusive, as Smith himself admits. The externalist claims that being motivated is not a necessary condition for being able to correctly use moral concepts, and, therefore, that an agent can make a genuine moral judgment without being motivated to act accordingly. On the other hand, the internalist claims that motivation is in fact necessary for moral concept mastery, and that the unmotivated agent is not *really* making a moral judgment, but simply uttering a moral proposition which she does not believe or mean. She is not a true amoralist because a necessary condition of being one (apart from being rational and unmotivated) is making actual moral judgments.⁴ Smith claims that the externalist cannot maintain his position without begging the question against the internalist, and that the dispute can only be settled by an independent account of what it means to master a concept.⁵

Still, consider this example: I am a deeply moral person who is always motivated to act according to her moral judgments, but, on one occasion when I judge that Φ -ing is right, while in conditions of full rationality, I do not feel motivated to Φ . In order to maintain the internalist thesis, I would have to accept that, even though I used in my judgment the same moral concepts that I have correctly used in the past, when (and, according to the internalist, because) I have been motivated, and even though I am in the same psychological state as on previous occasions when I have made a similar judgment, on this

²Ibid. p. 60.

³Ibid. pp. 66–67.

⁴Ibid. pp. 69–70.

⁵Ibid. p. 70.

occasion only I do not master the relevant moral concepts. Surely this is implausible. Without taking this to be a conclusive argument, I note that to call the person in this example an amoralist and to accept that genuine moral judgments can exist without motivation would not be question-begging, but would instead amount to an inference to the best explanation.⁶

Smith's second argument against externalism is more important. An externalist must appeal to a *de dicto* desire present in the agent in order to explain why a change in moral motivation usually follows a change in moral beliefs: namely, the desire to do what is right. But, Smith argues, this is a form of moral fetishism. Good people are not concerned with whatever happens to satisfy the predicate '___ is right' on a certain occasion, but instead care non-derivatively for their peers' wellbeing, for honesty, charity and the like.⁷ A better explanation for the 'reliable connection' between a change in judgment and motivation in a good and strong-willed person is that a belief that Φ -ing is right produces in the agent a direct desire to Φ , as a result of the Practicality Requirement.⁸

The externalist has two options open here. One is to simply deny that *de dicto* desires are necessary to explain the change in moral motivation. The good moral agent is simply someone who has a disposition to desire *de re* whatever she believes is right – no deriving is involved.⁹ A second option is to accept that *de dicto* desires to do the right thing do exist in such agents, but to deny that this is a form of fetishism. A good approach would be to claim that general *de dicto* desires give rise to appropriate *de re* ones – for example, it's not the case that I help the homeless man in the street because the homeless man is a prop in my attempt to do whatever is right, but rather because my desire to do the right thing has enabled me to feel direct concern for this man's welfare.

Accepting Smith's objection also entails accepting some counterintuitive consequences. For example, direct concern for honesty is a trait of the good person.¹⁰ But if derivativeness is a flaw, it follows that a good person is not allowed to care about honesty in general, but must spontaneously acquire a *de re* desire to be honest in every circumstance where honesty is required, without so

⁶See Brink [1] pp. 22–25.

⁷Smith [4] pp. 75–76.

⁸Ibid. 72–73; see also Section 2.

⁹Copp [2] pp. 50–51.

¹⁰Smith [4] p. 75.

much as using the concept of honesty in her moral deliberation, because concern for honesty as an abstract value entails a *de dicto* desire to do whatever the honest thing is. But this just contradicts our intuitions about morality. Moreover, the very idea of moral deliberation, especially in doubtful circumstances, requires a *de dicto* desire to do the right thing. Imagine a good, virtuous person who has to choose between several courses of action, some of which may be right, some of which may not be. If we accept that the purpose of moral deliberation is precisely to find *the* right action(s), then it is presupposed that the agent does not yet have a desire for the right action *per se*, since she does not yet know what the right action is, but that she has a *de dicto* desire to do whatever is right; and this is a trait of a virtuous agent. Ordinary moral practice and common sense suggest that *de dicto* desires are indispensable for virtuous agents and cannot be erased from any account of morality.

Although Smith's decisive objection to externalism – that deriving direct desires for the right action from a more general desire is a form of moral fetishism – fails, I ask the reader to keep it in mind because it will resurface soon.

2 Smith's rationalism and the need for externalism

This being said, let us examine Smith's own solution to the moral problem. This solution is based on the following analysis of rightness or normative reasons.

A1: Φ -ing is the right thing for agent *A* in circumstances *C* iff *A* would desire herself to Φ in *C* were she fully rational.¹¹

Although the analysis has been convincingly rejected,¹² I will assume it to be true in order to explore the consequences which follow from it, together with Smith's type of internalism and his acceptance of the Humean Theory of Motivation. Smith's moderate form of internalism is what he calls the Practicality Requirement.

PR: An agent who judges that Φ -ing is right is either motivated to Φ or is practically irrational.¹³

¹¹Smith [4] pp. 150–151.

¹²Copp [2], Brink [1], Sayre-McCord [3].

¹³Smith [4] p. 61.

And the Practicality Requirement is consistent with the Humean Theory of Motivation.¹⁴

HTM: An agent *A* is motivated to Φ in circumstances *C* if there is some Ψ such that the agent has a desire to Ψ and a means–end belief to the effect that Φ -ing would enable *A* to Ψ .

Combining A1, PR and HTM while presupposing moral objectivism enables Smith to attempt to solve the moral problem. By definition, the right thing is what a rational person would desire. So the rational agent will have an appropriate belief–desire pair (HTM) whenever she makes a moral judgment, meaning that a rational agent will necessarily be motivated by her moral judgments (PR) and will do so in virtue of the content of the judgment (A1), rather than in virtue of some disposition or desire she has.¹⁵

I reconstruct Smith’s argument for his rationalist internalism as follows:

1. If it is not the case that believing oneself to have a normative reason to Φ necessarily entails a corresponding motivating reason in a rational agent, then a rational person who makes a moral judgment, but is not motivated to act accordingly, is conceivable.
2. Anyone who believes she has a normative reason to Φ but does not have a motivating reason to Φ is necessarily irrational.
3. Therefore, the rational person will always be motivated (will desire) to act in accordance with her normative reasons.¹⁶

And premise two is supported by a further argument – a two-horned dilemma. Suppose I judge that it is right to Φ , but I am not motivated to Φ .

- 2a. I judge, correctly, that it is right to Φ ; therefore, I judge that if I were fully rational, I would be motivated to Φ . So, by my own admission, since I am not motivated to Φ at the moment, I am not fully rational; I am psychologically inconsistent.¹⁷
- 2b. I judge incorrectly that it is right to Φ but do not desire to Φ . I am irrational because I have the false belief that I have a normative reason

¹⁴Ibid. p. 92.

¹⁵Copp [2] pp. 36–37.

¹⁶Smith [4] pp. 177–179 and *passim*.

¹⁷Ibid. p. 177.

to Φ .¹⁸

It is quite easy to see that the second premise, on which the argument depends, is false. On the first horn of the dilemma which supports it, having the belief that Φ -ing is right in C does not entail having the belief that one would desire Φ in C if one were fully rational. Therefore, someone who has the former belief and lacks a desire to Φ is not irrational because there is no inconsistency among her beliefs, or between her beliefs and desires.¹⁹ Or an agent can recognise that she has a normative reason to Φ , but also that she has better normative reasons to Ψ , and therefore only develop a desire to Ψ and no desire to Φ .²⁰ Therefore, not desiring to do as one judges right is not irrational. On the second horn, Smith's account of rationality is implausible, because we do not usually equate rationality with omniscience and we allow rational people to have false beliefs or lack some relevant true beliefs.

Now, Smith claims that, to the extent that I am rational, I will acquire a desire to Φ as soon as I judge that Φ -ing is right, even if previously I have not had such a desire. The problem is that it is unclear how such a desire appears. Clearly, Smith's account demands that this desire be caused by the belief that Φ -ing is right, through moral deliberation.²¹ Smith talks about moral deliberation in terms of destroying old desires and acquiring new ones. Among the various ways in which we do this (including means-end reasoning and the use of imagination), Smith gives a primary position to the rational agent's attempt to make her desire-set more systematically justifiable, more unified and coherent and therefore rationally preferable. This includes (but is not limited to) acquiring desires that are consistent with her moral beliefs.²² Let us suppose, for the sake of the argument, that if I judge (correctly) that Φ -ing is right, acquiring a desire to Φ would indeed render my desire-set more rational, coherent, and unified than it is at the moment (but see Sayre-McCord [3] p. 75–76 for some plausible arguments that having a maximally coherent set of desires does not necessarily make one more rational).

¹⁸See *ibid.* p. 156, following Williams [5] p. 102 where having no false beliefs is a condition of full rationality.

¹⁹Copp [2] pp. 38–42.

²⁰*Ibid.* pp. 45–46.

²¹Smith [4] p. 179.

²²*Ibid.* pp. 159–160.

Smith talks as though acquiring desires is a simple process that can be triggered by the mere observation that it would be reasonable to have such-and-such desires.²³ In this case, the belief *B* with the content ‘ Φ -ing is right’ is supposed to cause in me a desire to Φ if I am rational. Leaving aside Smith’s framework, let us examine how this could happen. We are faced with two options. The first option is that my desire to Φ springs directly from my belief that Φ -ing would be right, with no interference from other pro-attitudes, dispositions or previous desires. The second is that *B* causes *D* thanks to a pre-existing motivational background which includes pro-attitudes and dispositions, such as a disposition to desire what I value. Now, in order to remain within an internalist framework, we must accept the first explanation. This is because the second explanation implies that *B* does not cause *D* by itself, but it simply activates a pre-existing motivational mechanism which then causes *D*. This implies that the same belief, if held by a rational agent who happens to lack such a mechanism (and Smith has not convincingly rejected the possibility of such an agent), is powerless in producing motivation; and this is externalism.

I think the first explanation is intuitively less plausible. Suppose I have the belief *B* with the content that Φ -ing would be right which rationally should cause in me a desire to Φ (*D*). Also suppose I have a strong pre-existing desire not to Φ (*D*′). Smith describes the process of acquiring *D* by analogy with the way in which we destroy and acquire new beliefs: if I believe that it is rational for me to have the belief *B*, then it is rational to acquire *B* and discard any beliefs which contradict it.²⁴ But this analogy does not take into account a crucial difference: when we derive a belief from another belief, we move from a mental state to another mental state of the same kind. But in our case, *B*, a mental state with a mind-to-world direction of fit, is required to directly cause *D*, a mental state with a world-to-mind direction of fit. Moreover, in the process, *B* must also overcome a very strong desire, *D*′, whose propositional content contradicts *B*’s own. And *D*′, given its direction of fit, requires that the world, of which *B* is a part, be changed in order to accommodate it. It seems to me entirely unclear and mysterious how this is possible completely independently of other background desires or pro-attitudes.

Smith could argue that, by definition, if I do not have the relevant desire, I am irrational and therefore fall outside the scope of PR, so my failure to be motivated does not count as evidence against internalism. But this only avoids

²³Ibid. pp. 159–160.

²⁴Smith [4] pp. 178–179.

the problem, since what we need is an explanation of how motivation appears when I *am* rational. First of all, the consideration that acquiring *D* would make me more rational is still just a descriptive observation and it remains unclear how it causes a pro-attitude, unless I have a background desire to be more rational. And, second, it is possible to imagine a rational agent who remains unmoved by this observation, e.g. because, even though she has a normative reason to Φ , there are other desires she can acquire instead which would render her desire-set even more rationally preferable.

Williams makes a similar point in noting that our ‘external’ (normative) reason statements, such as ‘it would be right to Φ ’, will not motivate us unless there already exists in our subjective motivational set (*S*) some element such that, when we deliberate from it, we become motivated to act as the external reason statement demands.²⁵ But, in this case, the external reason statement is superfluous, because an ‘internal’ (motivating) reason statement would have applied in the first place. For external reason statements to be genuine, rather than disguised internal reason statements, they must motivate us in the absence of any prior motivations we might have that could, through deliberation, lead us to Φ ; reason must be able to create motivation out of nothing, as it were.²⁶ Given Williams’s account of moral deliberation, this seems impossible, because we can never acquire new internal reasons – we can just discover internal reasons we have but are not aware of, or eliminate existing internal reasons based on false beliefs.²⁷ I take the latter claim to be implausible because it entails that people can never radically change their motivations, or, if someone does acquire motivations opposite to the ones she had before, it implies that she had had contradictory elements in her *S* all along. My view is that reason can cause any desires in an agent, as long as the agent has the right disposition or *de dicto* desire that enables her to desire whatever she values. Moreover, it is more economical to suppose that a single desire or disposition could enable virtually unlimited types of desires to be caused by corresponding beliefs than to suppose that, for every potential new motive, there must be some separate element in *S* that would lead to it through deliberation.

Smith could respond that all this is no threat to internalism because it is all presupposed by his account.²⁸ Internalism does not require, as I have claimed

²⁵Williams [5] pp. 107–108.

²⁶Ibid. p. 109.

²⁷Ibid. pp. 103–105.

²⁸See Williams [5] p. 109.

above, that moral beliefs cause desires in a vacuum. A rational agent is by definition someone so disposed as to acquire appropriate desires as a result of making moral judgments, and who also desires to become more rational. So, if I judge that Φ -ing is right, but I do not desire to Φ , I am not fully rational, but I am still rational enough to desire to make my desiderative profile more consistent and unified, and rational enough to be disposed to derive a *de re* desire to Φ from the belief that Φ -ing is right. This is consistent with PR, since it is not the agent's disposition that does the motivating, but the moral belief itself as long as it is held by the appropriate kind of agent.

But, to begin with, it is simply not true that the rational agent will necessarily be disposed to desire what she judges to be right. This only holds as long as we accept A1, i.e. the claim that what is right is what a rational agent would desire. But the argument above, which purports to prove A1 from premises 1–3, has been shown to fail, and, as pointed out through counterexamples, failing to desire as one judges right is not necessarily irrational.²⁹ The only way in which an agent can count as irrational while judging that Φ -ing is right but failing to desire to Φ is 'by her own lights', i.e. if she believes, while judging Φ to be right, that Φ -ing is what she would want to do if she were rational.³⁰ But many people do not think of rightness in these terms; and, for Smith's argument to be successful, everyone would have to have A1 in mind when making moral judgments.³¹ So, if we need to appeal to a certain disposition in the agent to explain changes in moral motivation, we must necessarily reject internalism.

And even if A1 were correct, I think Smith's brand of internalism would still lead us back to externalism. For, to begin with, the *de re* desires caused by beliefs would be derived – and doubly so, from both our moral judgments and from our disposition to be motivated by our moral judgment, which is exactly the objection that Smith brings against externalism. Even worse, everyone who judged that Φ -ing is right would, by definition, judge that Φ -ing would be the rational thing to do, and thus the desire to become more rational would play some role in moral motivation. Acquiring a desire to Φ because one desires to do the rational thing is fetishistic,³² by Smith's own criteria, and arguably even more so than desiring to Φ because one desires to do the right

²⁹See Sayre-McCord [3] pp. 74–76, Copp [2] p. 46, Brink [1] pp. 18–20.

³⁰Smith [4] p. 179.

³¹Sayre-McCord [3] p. 77.

³²See *ibid.* p. 76.

thing.

Analogous objections would apply, *mutatis mutandis*, to any analyses of rightness which follow the model proposed by Smith. Namely, any analysis which has the form

A': Φ -ing is the right thing for agent A in circumstances C iff A would desire herself to Φ in C were she ____ .

would demand the externalist condition that the agent be in a certain disposition, or have a certain *de dicto* desire, that enables her to desire *de re* whatever she judges to be right, as the best explanation as to how a pro-attitude can be caused by a purely descriptive mental state. Given all this, externalism seems like a simpler and better explanation of the phenomena that occur when motivation arises in a previously unmotivated agent. It has been profitable to focus on Smith's rationalist analysis not only because it provides a model for all other analyses, but also because, if it were correct, it would have an obvious advantage over them from an internalist perspective: explaining why an agent who judges that Φ -ing is right but is not motivated to Φ would appear to be irrational.³³

3 An alternative analysis of rightness

I take it to be a platitude that to say that Φ -ing is right is the same as saying that Φ *should* be done, or that *it would be good* to Φ , and consequently that there is a reason to Φ . To say that there is a reason to do Φ does not entail that not Φ -ing is irrational, for 'reasons' can be of many kinds, only a subsection of which includes reasons of rationality.³⁴ Saying that there is a reason to Φ also does not entail that this reason should be overriding. The point is that to judge that the world *should* be such-and-such is different from noting that the world *is* a certain way; it is simply part of the meaning of the world "should" that there is a possible state of the world and some reason to want this state of the world to be actual. We usually assume that whoever uses the world "should" will display this sort of pro-attitude herself, but this is not necessary; I can meaningfully use evaluative words without having any relevant feelings. But, plausibly, my use of the word "should" entails that I believe there is some reason for Φ -ing, even if it is not a reason I have; otherwise I would not have

³³See Copp [2] p. 38, Sayre-McCord [3] p. 67–68.

³⁴See Brink [1] p. 20.

bothered noting that there is some possible state of the world that it would be good to come about. But if *there* is a reason, it follows that, even if I do not have the reason to Φ now, I would plausibly have it if I were different – if I were the kind of person who fulfils whatever conditions are required of her to have a reason to Φ . An analysis of rightness needs to make reference to the concept of a reason, and there cannot be a reason without an agent to have a reason, and the agent making the judgment may or may not be this sort of agent. Hence the need for an analysis which is counterfactual in its formulation and makes reference to a kind of agent, such as A' .

No matter what we choose to fill the gap in A' with, as long as we accept HTM, we are required to also specify that the agent needs to have the appropriate dispositions or desires which enable her to desire whatever she judges right. So I propose we fill the gap with a property that entails this:

A2: Φ -ing is the right thing for agent A in circumstances C iff A would desire herself to Φ in C were she morally virtuous.

Now, A2 has almost no substantive content, but I think this is an advantage. We are not looking for a definition of what it is for an action to be right, but for a description of how we use moral concepts in deliberation. When I say that Φ -ing is right because Φ -ing is what I would do *if* I were morally virtuous, I remain non-committal as to whether Φ -ing is right *because* I would do it if I were virtuous, or, alternatively, whether Φ -ing is right independently of any agents and, were I morally virtuous, I would have no choice but to Φ . What A2 purports to do, as a non-reductive conceptual analysis, is to capture the ways in which we deliberate about moral matters. I believe that the openness of the concept of moral virtue, as well as A2's lack of commitment to either sort of causal relationship mentioned above, is an advantage in helping explain moral deliberation. We do not need to know exactly what makes a moral agent moral or what properties make a right action right; the point is that, when we engage in moral choices, we presuppose a certain conception of rightness and then we attempt to assess whether the relevant course(s) of action fit this conception. The openness of the concepts of rightness and moral virtue can explain moral disagreement and the diversity of outcomes in moral deliberation, since each of us presumably has a more-or-less fuzzy or changeable conception of the properties possessed by a morally virtuous agent.

Needless to say, further substantive analysis may show that the morally virtuous agent possesses other properties of the adequate kind, such as being rational, being compassionate, having certain desires, etc. But the main advantage of filling the gap in A' with 'morally virtuous', as opposed to any other property, is that it actually is a platitude that morally virtuous agents will desire

whatever is right,³⁵ while it is not a platitude that rational agents will do so, and the same applies to whatever other property we replace ‘rational’ with. As a result, the ‘morally virtuous’ predicate as applied to a moral agent entails, unlike other predicates, that the agent will be disposed to desire as she judges right.

It should now be clear how moral motivation appears and what happens if it fails to do so. Any agent could potentially make genuine moral judgments. If I judge that that Φ -ing is the right action in my circumstances, and I have the relevant disposition which enables me to desire to Φ , I will desire to Φ , regardless of any other pre-existing desires. I will be a morally virtuous agent on this occasion, and thus I will satisfy the condition specified in A2. If I simply lack the disposition which would enable me to desire to Φ , or my disposition is overridden by other factors, I am an amoralist, and I am this, to paraphrase Smith, by my own lights, since I judge that *if* I were moral I would desire to Φ . However, even if I am an amoralist on this occasion, my judgment that Φ -ing is right can still be genuine, as long as I master the relevant concepts. I may have a good image of what a moral person is, based on past experiences, my own previous moral choices, or people I know; the counterintuitive consequence of the internalist analysis, that moral people who are unmotivated on one occasion lose the mastery of moral concepts that they have displayed on previous occasions, is thus avoided.

4 Conclusion

Smith’s proposed analysis of rightness is insightful and useful, but insufficiently able to explain how moral agents derive a desire to Φ from the mere belief that Φ -ing is the right action. The natural explanation is to postulate some link between the moral belief and the desire it causes. But this is fatal to internalism. Firstly, this disposition is entirely independent of rationality; an agent can lack it and still be rational. Therefore Smith’s brand of internalism, PR, which claims that a rational agent will necessarily desire what she judges to be right, collapses. And if PR were true, it would still face the same objection based on derivativeness and moral fetishism that Smith brings against externalism. Externalism is then to be preferred to internalism, through an inference to the best explanation. This needs to be captured by an analysis of moral rightness which accommodates externalism, and I have attempted to

³⁵See Brink [1] p. 14–15.

provide just such an analysis.

References

- [1] Brink, D. (1997) ‘Moral Motivation’, in *Ethics*, 108: 4–32.
- [2] Copp, D. (1997) ‘Belief, Reason and Motivation: Michael Smith’s “The Moral Problem”’, in *Ethics*, 108: 33–54.
- [3] Sayre-McCord, G. (1997) ‘The Metaethical Problem’, in *Ethics*, 108: 55–83.
- [4] Smith, M. (1995) *The Moral Problem*. Oxford: Blackwell.
- [5] Williams, B. (1981) ‘Internal and External Reasons’, in B. Williams *Moral Luck*. Cambridge: Cambridge University Press.

On the Moral Status of Objectification in Pornography

Carl-Otto Frietsch

University of Warwick

Introduction

The aim of this essay is to argue that objectification is not a morally pernicious feature of pornography. This does not mean that it seeks to establish that such a thing is somehow *impossible*, but rather that a multitude of the arguments and theories advocating this view are misguided. Indeed, contrary to many commentators, I hold that not even instances of *sexual* objectification are able to capture what is wrong with pornography. By drawing upon Jean-Paul Sartre's work on autonomy and objectification in his book *Being and Nothingness*, I introduce a distinction between the *act of objectification* and *attitudes towards objectified individuals*. I hold that only the attitudes – which correspond to *how* one treats someone as an object, as opposed to the mere act of *making someone into one* – are morally objectionable. On this basis, I argue that Martha Nussbaum and Rae Langton are wrong to hold that objectification constitutes a morally reprehensible aspect of pornography because they fail to distinguish between attitudes towards objectified people and the act of objectification itself. With this distinction in place, I conclude that objectification is revealed as an integral part of everyday life that is unfit to bear the responsibility for the vices of pornography.

I will begin by providing an account of what Nussbaum and Langton take objectification and its relation to pornography to be. After doing so, I shall outline some of the key aspects of Sartre's conception of autonomy and objectification. I then intend to show how this reveals the distinction between attitudes towards objectified individuals and the act of objectification. I will further argue that the way in which this distinction excludes the act of objectification from moral considerations serves to solve several ambiguities present within Nussbaum and Langton's accounts, which is a reason why they should accept it. This is taken to justify the distinction, which further entails my conclusion.

1 Objectification in Pornography

In her book *Feminism Unmodified*, Catherine MacKinnon defines pornography in the following way:

We define pornography as the graphic sexually explicit subordination of women through pictures or words that also includes women dehumanized as sexual objects, things, or commodities; enjoying pain or humiliation or rape; being tied up, cut up, mutilated, bruised or physically hurt; in postures of sexual submission or servility or display; reduced to body parts, penetrated by objects or animals, or presented in scenarios of degradation, injury, torture; shown as filthy or inferior; bleeding, bruised, or hurt in a context that makes these conditions sexual.¹

The standard account of objectification is a constitutive aspect of this definition. The idea is that someone is appropriated as a thing or commodity to be used in a particular way by someone else. This popular way of understanding the term is highly influenced by Kant's formula of humanity, which holds that one should always treat other people as ends and never as a means.² When portrayed in this manner, objectification appears abhorrent because of a widely shared intuition that human beings are fundamentally *not* objects analogous to tools. MacKinnon's definition asserts that objectification is inherent to pornography and is thus responsible for a significant amount of women's suffering. Even if one is unwilling to accept this definition to the letter, one might still think that the general connection it makes between objectification and pornography is valid. In a paper called 'Objectification' by Martha Nussbaum, MacKinnon's connection plays a foundational role in relating these concepts to each other. However, the project of Nussbaum's paper is to make their relation clearer by providing a more detailed account of what objectification is.³ Summarily, she holds that an objectified individual may lack autonomy, subjectivity, be inert, or be used instrumentally as a tool of exchange, destruction or possession. This cluster concept is then not just used to explain objectification, but *sexual* objectification as well. On Nussbaum's view,⁴ women become subject to a case of sexual objectification when they are

¹ MacKinnon [4] p. 176.

² Kant [1] p. 176.

³ For her list and descriptions of these features, see Nussbaum [3] p. 257.

⁴ Nussbaum [3] p. 280.

robbed of their humanity by having their inertness and lack of autonomy used as an erotic resource for men. Accordingly, it is easy to see why Nussbaum holds that pornography embodies several varieties of sexual objectification and further thinks of these as being morally objectionable.

It seems reasonable to agree with Nussbaum that at least *some* parts of her cluster concept constitute intrinsic aspects of objectification – especially inertness, lack of autonomy and subjectivity. However, is this also true for instrumentality, exchangeability, violability and ownership? I deny this. I hold that these features are *attitudes expressed towards objectified individuals*, which must be sharply distinguished from the *act of objectification* itself. Furthermore, I maintain that it is attitudes like these – and not the mere act of making somebody into an object – that are morally pernicious. To reach this conclusion, the distinction between an act of objectification and an attitude towards an objectified individual must be argued for. I begin this argument by agreeing with Nussbaum that one of the core premises of her theory is correct, which is that objectification involves a denial of autonomy. In order to establish this, however, the concept of ‘autonomy’ must be explained more carefully than it is in Nussbaum’s paper, or else it might lead to confusion. In a response to Nussbaum, Rae Langton exposes this problem by arguing that there are *autonomy-affirming* kinds of objectification.⁵ Specifically, Langton holds that there are instances where one must affirm someone’s autonomy in order to deny it in such a way that objectification becomes possible. This seems *prima facie* paradoxical, but it is not. To illustrate her point, Langton considers a case of sadistic rape.⁶ A sadistic rapist does not simply wish to rape the victim, but wants the victim to *struggle* and *resist* whilst the act is committed. Langton holds that this requires an attribution of autonomy to the victim, which is thus integral to the autonomy-denial and objectification exercised by the rapist. She further maintains that this autonomy-affirming objectification constitutes a significant morally pernicious aspect of pornography.

It is clear that Langton’s perplexing pseudo-paradox arises because neither she nor Nussbaum spends much time elucidating and disambiguating the notion of autonomy. This, therefore, seems to be the next step in providing a satisfactory account of what objectification is. Once this has been done, I think that it will be possible to see why it is distinct from attitudes expressed towards objectified individuals, as well as why it is not a morally reprehensible element

⁵ Langton [2] p. 224.

⁶ Ibid. p. 234.

of pornography.

2 Autonomy and Objectification

The discussion between Nussbaum and Langton appears to generally understand autonomy in terms of – as Langton puts it – “treating somebody [...] as being capable of choice”.⁷ This seems to be a move in the right direction, but a more detailed analysis is required in order to provide clarity. This presents an opportunity to incorporate and appreciate some of the literature that deals extensively with autonomy and objectification but which has, despite this, been largely neglected by the contemporary feminist debate.

In his famous book *Being and Nothingness*, Jean-Paul Sartre argues that the autonomy of a human being is not given to her in virtue of her *being treated by others* as capable of choosing her own actions, but rather by *her awareness of herself* as a being that always has to choose her own actions.⁸ According to Sartre, every act that we recognise as our own is experienced as ultimately caused by our own choices, rather than by some force external to us. When I lift my hand, I am aware of this as an act that I *choose* to do, as opposed to something that simply *happens* to me (such as when I am pushed). Since one cannot choose *not* to choose, we thus live in a state of being aware of that we always have to choose our own actions insofar as we are conscious. There are ultimately no overriding external means capable of making these choices for us, nor relieving us from being responsible for them. On Sartre’s view this is what it is like to be autonomous – one is *condemned* to be free by continuously having to choose what to do next.⁹ He further claims that human beings systematically attempt to *escape* this autonomy by denying themselves the terrifying and overwhelming freedom it offers them. Sartre says that this involves an apprehension of oneself and others as *things*, which thus constitutes the *act of objectification*.¹⁰ An object is an object because it is not free, which means that it does not have to constantly burden itself with making choices. Sartre thus holds that autonomy-denial is the central aspect of all objectifying acts. For instance, an act of self-objectification attempts to relieve oneself of one’s freedom by pretending that one – like an object – does not have a choice. I

⁷ Ibid. p. 226.

⁸ Sartre [6] p. 75.

⁹ Ibid. p. 462.

¹⁰ Ibid. p. 67.

might claim that I am too addicted to smoking to quit. I thereby relieve myself of the responsibility of smoking by attributing my own choice to do so to some circumstance that I have no control over (my addiction). By doing so I make myself into a smoking-thing – an object whose essence is to smoke.

Of course, this is just an elaborate kind of self-deception. We are always aware of *not* being the objects that we all-too-often portray ourselves as being, as well as of *not* possessing any essential features that are able to externally guide our choices. This remains true even when we objectify *others*, which we do by denying that other people are always free to choose in exactly the same way as we are. For example, suppose I find myself speaking to a convicted rapist. Regardless of how many times he assures me of that he has changed his way of life and is no longer a predator, he still appears before me with his freedom limited in a way that makes him different from me. I still define him as a rapist, which restricts his freedom by fixing the pre-determined function of ‘someone who chooses to rape’ onto him. In the same way, other people might objectify *us*. Suppose I am caught shoplifting and therefore become labeled a ‘thief’ by the person who catches me. Here, I become aware of that someone else has made theft a part of me, even though I am painfully aware of that I have no such essential feature, since I am always aware of myself as capable of acting differently. By expanding upon this idea, Sartre holds that objectification actually constitutes an integral part of human relationships. Along these lines, it becomes possible to see why one might not view it as something that is morally wrong. For instance, a friendship might very well hinge on one’s conception of one’s friend as somebody who completely lacks the freedom to commit a homicide. Indeed, one of his virtues might be that one has completely robbed him of the choice to do such a gruesome thing. Accordingly, it appears very difficult *not* to objectify other people in some way. It simply does not seem possible to have access to the autonomy of others in the same way that one has access to one’s own.

This might seem depressing at first, but to feel that way is unnecessary. After all, even if we always objectify other people in some way, we can still *treat them* as if they were just as autonomous as us. Despite the fact that we cannot experience the autonomy of others, we can still recognise and respect them as ‘not adequately objectifiable’ beings. This is entirely dependent upon our *attitude* towards them.¹¹ I may not be able to stop myself from seeing the convicted rapist as a rapist, but I can recognise that it is always *possible* for

¹¹ This interpretation is highly influenced by Peter Poellner’s discussion of Sartrean existentialist ethics (Poellner [5] p. 19).

him to choose his conduct differently and that his crimes are therefore not a fundamental part of him. This marks the distinction between the *act of objectification* and *an attitude towards an objectified individual*.

With this important distinction out in the open, it becomes possible to see that Nussbaum's cluster concept is indeed *too* clustered. It confuses several *attitudes* – i.e. instrumentality, exchangeability, violability and ownership – pertaining to *how the objectified individual is treated*, with the objectifying act itself. It further seems that the moral connotations that Nussbaum wishes to attach to the act of objectification are more effectively accounted for if explained by a multiplicity of such attitudes. This is because her account concedes that there are several elusive instances of objectification that are not morally objectionable at all alongside the ones that are. She even holds that some instances of sexual objectification may constitute a wonderful part of one's sexual life.¹² I believe that this may be explained through an objectifying act that is accompanied by an attitude that recognises that the objectified person is actually not an object, but rather *just as autonomous as the objectifying person himself*. This does not appear to be morally reprehensible in the slightest. Indeed, it seems to be exactly what is happening in the cases of sexually pleasing 'surrender' that Nussbaum talks about. On the other hand, the occasional moral wrongness that she also associates with objectification and sexual objectification may be explained by the presence of an attitude towards the objectified person that denies the autonomy of that person. Langton's pseudo-paradox – now explainable in a less ambiguous way – further seems to constitute such a case. It is obviously not true that the sadistic rapist attributes autonomy to his victim. Apart from objectifying her, he treats her with an instrumental attitude that denies her autonomy. This treatment is a denial of her autonomy due to its utter incompatibility with the recognition that she is *just as autonomous* as the rapist is, which is the present necessary and sufficient condition for attributing autonomy to another person.

What the rapist does relies on is the victim's struggling and resisting, *but only to the extent* that the rapist himself finds it sexually arousing. Contrary to attributing autonomy to her, the whole act presupposes that she is treated as an object that *only* makes the choice to struggle and resist in the right kind of way. The same is also true for Langton's remarks on the pornographic film *Deep Throat*, which, she holds, is another example of when women are given 'autonomy' so that it may be taken away from them. The film features a 'sex-

¹² Nussbaum [3] p. 251.

ually liberated' female protagonist who "[...] is driven by nothing but an insatiable desire for throat sex".¹³ The film claims to attribute autonomy to the protagonist by emphasising her freedom to choose, but the present analysis reveals this to be false. The 'autonomy' is attributed to her by a continuous choice to *have more throat sex*, rather than by the fact of her own choosing. Accordingly, *Deep Throat* promotes and perpetuates an attitude towards women that holds their autonomy to be realised by their always making a *particular choice*, as opposed to by their choosing in general. This finally demonstrates how an attitude towards an objectified individual may constitute a morally bad element of pornography. By contrast, the act of objectification figures as a constitutive element of human relations that is too general to acquire a specific moral status – even when those relations are sexual. On this basis, the attitudes relating to how one treats objectified people seem to be much better candidates for explaining what is morally wrong with pornography than the mere act of objectification is.

3 Conclusion

In conclusion, I hope to have provided a convincing account of why objectification is not a morally pernicious feature of pornography. Contrary to what Nussbaum and Langton suppose, the moral aspect of the relation between pornography and objectification resides with the attitude expressed towards the objectified person, rather than with the objectifying act itself. What I have sought to show is ultimately that objectification plays a much more fundamental role in our society than the contemporary feminist debate acknowledges. A careful examination of the concept reveals that it is incredibly difficult to supply an account that renders it either decisively good or bad. I believe that Langton and Nussbaum's analyses benefit from this Sartrean insight, for which reason I have tried my best to bridge whatever 'divide' there is between Sartre's account and theirs.

References

- [1] Kant, I. (1996) *Practical Philosophy (The Cambridge Edition of the Works of Immanuel Kant)*. Mary J. Gregor (ed. & trans.). Cambridge Cambridge

¹³ Langton [2] p. 237.

University Press.

- [2] Langton, R. (2009) *Sexual Solipsism: Philosophical Essays on Pornography and Objectification*. Oxford: Oxford University Press.
- [3] Nussbaum, M. (1995) 'Objectification', in *Philosophy & Public Affairs*, 24: 249–291.
- [4] MacKinnon, C. (1987) *Feminism Unmodified*. Cambridge, Mass: Harvard University Press.
- [5] Poellner, P. (2012) 'Early Sartre on Freedom and Ethics', in *European Journal of Philosophy*, 24: 249–291. Published online (17 APR 2012) at EJP Early View: <http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0378.2012.00532.x/abstract>
- [6] Sartre, J.P. (2003) *Being and Nothingness*, trans. H. Barnes. London and New York: Routledge.

Can We Practically Measure Freedom?

Alexander Tolcher

University of Exeter

Assuming that freedom is measurable, why might a measure of freedom be useful? Well, reflection seems to provide several obvious uses for such a measure. For example, if we accept that freedom in itself is valuable, either for instrumental reasons or for its intrinsic value, then being able to assess our actions (or the actions of a group) in terms of their effects on freedom would be valuable in helping to analyse the benefits of existing political systems. In what follows, I will assume that freedom is valuable in this way and that freedom is to be thought of in a negative sense: as an opportunity concept as opposed to an exercise concept and as a non-preference-based concept.

This potential valuable use for a measure of freedom relies on the measure's being applicable to real-world situations. However, most of the literature is silent on how application might be achieved. As Graeff writes, there is "a gap between theory and empirical operationalization".¹ This seems to deprive the topic of measuring freedom with a value that it could possess, a fact that is all the more worrying given that some writers like Pattaniak and Xu justify their measures in similar terms to the ones outlined above without giving any account of application at all.²

The aim of this essay is to solve this problem by attempting to find a model for how one might practically operationalise understandings of negative freedom. I will survey the most prominent existing types of negative freedom measures and show that none of them are applicable as they stand. Next, two practical methods of measuring freedom will be discussed both as solutions to the problems encountered by theoretical measures and as stand-alone practical measure: first, an 'impressionistic' model of measuring freedom based on a suggestion in the work of Isaiah Berlin; second, several 'indications'-based models, used in various forms both in the literature and in institutionally attempted measures. It will be shown that none of these can stand either as reliable indicators for any of the theoretical measure surveyed or as practical measures in their own right. In the last part of the essay I will argue that there is a way to solve the

¹Graeff [3] pp. 113–132.

²Pattanaik and Xu [6] pp. 383–390.

problem of practically measuring freedom. This solution will take the form of a workable model of application for both the cardinality measure and the social measure of freedom rather than an independent measure. This is for two reasons: firstly, given that it seems that these theoretical measures can be made applicable there is no real need to create a new measure entirely; and, secondly, the measures proposed are good attempts to philosophically operationalise the concept of negative freedom in its theoretical form, so it seems wise to make use of such work and to work to apply it.

The first step required in a search for a solution to the problem of inapplicability faced by current philosophical measures is to make sure the problem really exists. This is not shown by the absence of discussion of the issue of applicability in the literature. To show that such a model is required I shall now look at three common, non-preference-based measures and show that, as they stand, they are not applicable.

Consider first of all the measure of cardinal freedom as put forward by Patanaik and Xu.³ This measure calculates the freedom of the individual by summing together all the relevant freedoms which that individual possesses. However, applying this measure is obviously problematic. It seems far from obvious how we are to work out the number of freedoms open to an agent. Indeed, on reflection, such an exercise appears impossible. This is evident as if one accepts that the freedom of an agent must be the freedom of an agent to act in some way and proceeds to try to count the actions open to the agent, as on any common principle of act individuation the number of actions open to any individual (even the most lamentably unfree) would have the appearance of infinity.⁴ Therefore, the activity of enumerating and identifying this number of actions, even approximately, would be impossible without a model by which to provide a reasonable approximation for the sum of these acts (hence the need for a model of application). This problem of practical enumeration

³Ibid.

⁴To see this clearly, consider the issue with regard to two popular models of act individuation. (1) Individuation by physical properties: The number of different combinations of physical properties can be possessed by object *X* is infinitely vast if one considers the size of the smallest increments in any physical (including temporal) dimension and then considers on that basis how many physical property ascriptions could be fulfilled simply by the ways we can wiggle our thumbs. (2) Individuation by time, agent concerned and act attribute: Given that act attribute on this theory can include any possible attribute that can be attached to an action, even if all that was allowed were common language ascriptions, the number of ways in which one could describe any given action would be vast given the many possible precision-increasing additions of adjectives to actions and the many synonyms possible for every attribute ascription.

of relevant factors in the measurement of freedom I shall call the ‘problem of counting’. This problem is equally applicable to Carter and Steiner’s social measure of freedom as this, too, requires summing all the actions (or choices) an agent is free to perform and, in addition to this, requires further that this number be divided through the number of actions that are technically feasible for that agent.⁵

The diversity measures of freedom such as those of Suppes and Rosenbaum also suffer from the problem of counting when it comes to empirical application. However, this approach does not suffer in precisely the same way.^{6,7} The measure attempts to work out individual freedom based on the sum of the greatest differences between all the characteristics of the actions in a particular choice set (or, in Suppes case, the entropy of the set in question). This falls foul of the problem of counting in the following respects. Firstly, in order to be sure that attributes *X* and *Y* were separated by the greatest distance on a linear attribute scale, one would have to know at least most (ideally all) the positions of the other points in order to be sure of the place of *X* and *Y*. Secondly, if one were trying to measure the freedom of a broad set, where the relevant characteristics, whose maximal variation should be taken into consideration when measuring the total maximal variation of that set (i.e. the freedom of the set), are obviously great (for example in the case of overall freedom), then one would have to enumerate all of these characteristics of variation. In the case of overall freedom, all possible characteristics of variation are necessary, and enumerating these is obviously as impractical as enumerating all the different actions one is free to do. Thus, the diversity measure also falls foul of the problem of counting and so cannot be applied to the real world as it is.

It has been shown in the first part of this essay that all major approaches to the measurement of negative opportunity-based freedom are inapplicable as they stand. This next part of the essay will consider various models for practically estimating freedom to see if any might be used to apply the above theories (either by providing a reliable indicator of what the result of the above theories would be when applied, or by providing a reliable indicator of the variables necessary to apply them).

⁵The fact that these two measures fail in terms of empirical applicability obviously entails the failure of hybrid forms which incorporate both measures, such as the measure Kramer introduces in *The Quality of Freedom*.

⁶Rosenbaum [7] pp. 205–227.

⁷Suppes [9] pp. 183–200.

A possible solution to this problem, as well as a possible independent measure, is suggested by Isaiah Berlin's comment in a footnote in his essay 'Two Concepts of Liberty'. Here, he states that "the method of counting these [possibilities of action] can never be more than impressionistic".⁸ This suggests one possible way of estimating the variables of the above approaches and, thus, a possible model of application as well as a possible independent measure that could stand as a practical measure of freedom.⁹ Unfortunately, it fails in both these respects. Firstly, impressions will be highly contingent; for example, impressions of the possibilities of action opened up by various aspects of society (e.g. a democratic government) are likely to vary between individuals based on their personal experience and cultural background. Equally, they would also likely result in error due to subjective bias, as, regardless of the amount of data each participant was presented with, it is likely that they would naturally take greater account of the freedoms (or the elements of a choice set) which were 'significant' and may fail to take account of various 'insignificant' freedoms not widely (if ever) discussed in society generally. For example, many people would not take into account the large number of actions from which one is physically barred (by law enforcement) from carrying out, in the case of legislation, which only allows people to drive on one side of the road. The judgements of those taking the impressions would thus likely be affected by factors apart from the actual freedoms present in an overall freedom set. A better model should therefore be sought.

There are several examples of more systematic attempts to practically measure freedom. All these attempts can be described as measuring freedom through finding measurable indications of it (hence I shall call them 'indication' measures of freedom). I will survey this type of approach in order to find whether any of the models proposed could be used to measure freedom on their own, or to help to apply the measures given in the first part of the essay. Given that we are trying to find a practical measure of overall freedom, I will only focus on the indication-based models which claim to indicate overall freedom.

One significant indicator model present in the literature is Carter's joint freedom metric.^{10 11} The first part of this metric is simply a non-moralised list of

⁸Berlin [1] p. 128.

⁹Note that this model would not be trying to take into account a subject's feeling of freedom or their impression of freedom itself but their impressions of certain variables.

¹⁰Carter [2] pp. 269–288.

¹¹This metric attempts only to count the extent of freedom and could not provide information on

general freedoms which are all indicators of the presence of a number of possible actions (hence indicative of Carter's theoretical measure) based on which each country is to be given a score. The extent of individual freedoms that that freedom covers, the extent to which that freedom covers actions also covered by another and the extent to which the listed freedom are restricted (taking into account threats as reliable indicators of a restriction of relevant compossible actions). The second part is his exchange value metric, which only provides information on the amount and number of goods an agent might exchange what he has for (based on his possession of exchange value). However, not only is this very limited but it is radically dependent on other freedoms, such as freedom of exchange, freedom of enterprise and so on; without these freedoms, having goods by itself has absolutely no impact on freedom. Therefore, Carter's second measure only works if one assumes that these are present and equal in all societies, which they are not. Even if this is not enough to dismiss the second metric all together, it can at least be concluded from the above that it is radically incomplete as a practical operationalisation of freedom if the first metric fails. And it is evident that the first part fails on the following considerations. It requires the selection of relevant freedoms, which are to form the basis for the indication of freedom, without having any method of selection. Without a selection process, the selection of freedoms will be based on individual (or on group) judgements as to which freedoms to include. In this selection it seems probable that their judgments will be biased according to contingent factors, such as presence or absence of each freedom in their cultural discourse and personal histories. Without a stringent method for selecting a freedoms list, these biases seem extremely likely. (What other means are there for selecting freedoms to go on the list other than the contingent factors mentioned?) If these contingent factors form the basis of the metric, then the metric cannot adequately indicate any of the theoretical measures previously mentioned, as different applications in different circumstances will yield different results with no way of deciding between them. Also, with this contingent basis, the measure would be affected by other variables besides overall freedom, which would result in the measure having a serious endogeneity bias, and would thus be unreliable.¹² For these reasons, the metric cannot be taken to give a reliable indication of the actions one is free to perform and so is of no help in applying any of the theoretical measures of freedom (which was Carter's aim). Nor, for

the extreme elements in freedom sets. For this reason it could only be used to apply social and cardinal measures.

¹²While it could be argued that cross-cultural consensus could solve this problem, it would be an empirical question as to whether a sufficient consensus could be established.

the same reasons, is it a good independent practical measure of freedom.¹³

Another significant indicator model uses the rights (and/or civil liberties) an agent possesses in order to indicate his freedom. There are two versions of this model. The first, which forms part of the measure of total freedom conducted by the State of World Liberty Index, as well as other NGOs in the past, creates a checklist of rights and then rates each country or individual based on that pre-prepared checklist. A problem with many of these lists, as Carter has pointed out in reference to some early ones, is that they tend to be moralistic (as it is with the HFI) and so would not represent or even take into account non-moral (or even amoral) freedoms therefore would not provide a sufficient model for indicating any of the theoretical measures.¹⁴ Although the model could be altered so that the ‘rights’ listed took into account non-moral (and amoral) freedoms, it would, without a stringent system for deciding what to select and without giving some assurance that the list reflected overall freedom, fall into the same problem as that encountered by Carter and discussed above. Therefore, this version of the rights indicator model should not be accepted as a reliable indicator of overall freedom or as a means of examining representing the free action of an individual or group in order to apply the theoretical measures above.

Another slightly different approach is to take into account the extent of the rights an individual actually has without coming at the inquiry with a checklist. This is endorsed by Graeff and rests on the assumption that any rights (properly enforced) do entail a space of social freedom.¹⁵ While this looks like a sound basis for overall freedom, there are a number of problems with Graeff’s measure. Firstly, Graeff himself points out that many of the rights that legitimately entail freedom might overlap to some extent so that the amount of freedom they represent is greater in some samples than others and some rights do not entail much freedom at all. For example, if we take into account rights in a broad sense (such as would be amenable to empirical testing), then the right to democratic government may be affected both by the freedom to vote and the governmental structure (which would not necessarily imply greater freedoms). Secondly, it is possible that there may be *de facto* freedoms (or

¹³Even if the problem of contingency in Carter’s measure could be solved, the metric would still face a greater one as the list of freedoms could never hope to be even vaguely comprehensive, and so it is seriously doubtful (but not impossible) that any list of freedoms could adequately represent or be used to indicate the overall number of freedoms of an individual or group.

¹⁴Carter [2] p. 276.

¹⁵Graeff [3] p. 128.

unfreedoms) not elucidated by the rights that people have. Lastly, in order to use rights data to take into account freedom of action, it is necessary to take into account the extent of right enforcement. For example, Oppenheim asks how we are to compare freedom from spending fines between countries where the enforcement differs; answers to these type of questions are absent from the rights measure.¹⁶ For the above reasons, this model cannot be seen as a reliable indicator for any of the theoretical measures proposed in the first part of the essay, as it does not reliably represent the number or type of free actions. Further the same failings means that it cannot be seen as a reliable independent practical measure.

In this last part of this essay, I will attempt to sketch a model which aims at supplying a reliable model for applying (via reliable indication) the social measure of freedom and the cardinality measure (and so by extension Kramer's measure also); but not the diversity measure.

The first part of the measure takes into account freedom of movement. This part takes as its inspiration Carter's spacetime grid as proposed in 'A Measure of Freedom'.

Space and time must be thought of as an immobile *grid*, made up of a finite series of spatio-temporal *regions*. A physical object of a standard volume can then be seen as a potential occupant of one or another [...] fixed region.¹⁷

This suggests thinking of actions as involving certain pieces of matter, each of which has the opportunity to occupy a finite number of spaces at a finite number of times. The extensiveness of the freedom of an action depends on how many spacetime regions each of the set volumes of matter can occupy. This idea is meant to provide a solution to the problem of accounting for the differing extensive properties of different freedoms of action. However, I propose an alternate use for this freedom grid. I propose that we use it as part of an attempt to find a way to practically operationalise freedom of movement as the first part of an attempt to find a practical measure for the freedoms of an individual or group which can be used to apply the cardinal and social measures of freedom.

This part of the measure is as follows. Divide the world's surface into equally sized areas. These need to be fairly small in order to make the model accurate.

¹⁶Oppenheim [5] p. 184.

¹⁷Carter [2] p. 184.

However, this will obviously be limited for practical concerns. Take an individual (or a sample of multiple individuals in the group one is trying to investigate) and for them (or each of them) shade in the area of the world in which they cannot go (this will generally involve shading in the countries which the individual is not free to go to but will also involve particular places). Given that where someone can go often depends on their status, gender, nationality or race, and the places to which these attributes prohibit going are usually well known, this part will be relatively easy. This information would enable the deployment of the social and cardinal measures of freedom with regard to this freedom of movement. Quite simply, the areas where one is able to go would be counted to give the cardinal measure and this could be divided over itself plus the total area where one is not free to go to give the social measure. This would take into account freedom of movement.

This is not the only freedom we should account for. Therefore, I propose to add freedom of action into this model. In order to do this divide the surface area in which one is free to travel into ‘areas of general homogeneity’ in which there is little or no difference between law, social and cultural attitude, socio-economic circumstance and the sanctity of that area. For each of these homogeneous zones take a random sample of areas and within each area work out in detail the number of actions the individual is unable to do which are available as possibilities elsewhere in the world. The reason for this approach is that those actions which one is unable to do will be far more obvious than those which are available and could be reliably ascertained by understanding the social and legal conventions; generally, prohibitions tend to be far more obvious than licences. The focus on particular areas means one would be better able to get a sufficient understanding of the social and cultural factors involved through interviews and even be able to use trial documents in order to work out how the law is practised and interpreted in the region. If we disallow the possibility of freedom’s being afforded by technology (or any other form of positive enabling), as is consistent with a measure of negative freedom, and assume that unfreedom is inversely proportional to freedom, then one would now be in a position to deploy both the cardinal and the social measures of freedom of freedom for the randomly sampled areas.¹⁸ If freedom is inversely proportional to unfreedom then, as unfreedom increases, cardinal freedom would be reduced. Then all one would have to do would be to posit an arbi-

¹⁸Not an unreasonable assumption as, if one discounts technology, freedom would be of a Hobbesian sort, i.e. without limitation or coercion, thus any unfreedom would mean a step away from this total freedom and thus less freedom; and therefore the more unfreedoms the less freedom.

rary large number for the largest freedom of action possible and subtract the number of unfreedoms from it to create a measure informed by the cardinal measure, and which would yield similar comparisons to the ones that would be hypothetically afforded if the cardinal measure were used in its theoretical form. To deploy the social measure one would simply have to take the number of unfreedoms and divide it by itself and the number chosen as representing the maximum number of freedoms. Although guessing the maximum number of freedoms may seem arbitrary and so seem to defeat the point of this approach, this is not the case. Whatever the number chosen to represent the maximum number of freedoms, the rankings, which would be produced by both measures (e.g. the ranking of one person as more free than another), would remain constant. Since the ability to make empirically grounded, reliable comparisons between people and groups is the main aim of the measures of freedom, this form of operationalisation serves its function well. However, it must be said that this does create some difficulty for adding freedom of action to the general freedom of movement, though the problem (which is addressed below) is not as bad as it first appears. Using what has been said above, we would be able to work out the freedom, in terms of the social and cardinal measure, of the randomly sampled areas within the areas of general homogeneity.¹⁹

At this point we have a series of random samples in a generally homogeneous area. In order to give a representation of the overall freedom of action an average measure will be created for each area of general homogeneity by averaging all the random samples from that area. Next, each average figure for each area will be multiplied by the fraction of the total area of free movement it cov-

¹⁹If one did not wish to make the above assumptions and include technology then, to the measure of unfreedoms, one could start by either creating a list of general technology types and then measuring then number which the sample would be unable to get hold of (due to money, position, culture, social pressure, etc.), or one could use company product portfolios for all registered companies and perform the same operation of adding those unavailable to the value of the measured unfreedoms. With this figure, the operations described above could be made in order to approximate the data necessary to deploy the social and cardinal measures of freedom. However, there are good reasons to avoid this. Firstly, there is the general opinion, voiced by Isaiah Berlin (amongst others), that freedom and ability are two different things. As Helvetius puts it, "It is not lack of freedom not to fly like an eagle or swim like a whale." Secondly, neither method measured above seems to offer results that are commensurable with the values accorded to the literal unfreedoms, as they do not describe actions but the availability of technologies. But surely counting the actions you could and could not do with a technology would be impossible. There is no data from which one can get even an approximate idea of the number of actions involved in having or not having technology; this is unlike unfreedoms resulting from constraints, as the action (or actions) one is not permitted to do are usually explicitly given. Therefore, on both philosophical and practical grounds, it would be best to exclude this; I have included it as an option for the sake of completeness.

ers. The resultant figures will then be summed to create the average amount of freedom the individual has in the areas in which he is able to move. This figure can represent freedom of movement per square. Now, it is obvious that, if 5 is the average number of free actions per square (or 0.5 the freedom of action in terms of the social measure of freedom), then the more squares one is free to occupy the more free actions one is able to perform. Hence, the average figure per square should be multiplied by the number of squares the individual is free to occupy. Thus, we have a measure of freedom of action which does not rely on subjectively chosen impressions or conditions but, instead, takes into account the individuals' (or groups') situation in its totality in a way which is capable of operationalising the cardinality measure and the social measure in terms of freedom of action. It might seem that this would render obsolete the 'freedom of movement' measure worked out earlier and, indeed, either can be used independently and both function as a practical operationalisation of the cardinal and social measures. However, it seems that the ability to go to one place or another is a freedom separate from freedom of action, which could not be accounted for in the microanalysis which forms part of our freedom of action calculations. Therefore, I propose we try and combine the two.

Here it must again be stressed that the aim of this exercise is to avoid arbitrary measurement in order to create a non-arbitrary freedom ranking of people and groups. The exercise is not meant to create the exact figure that the theoretical measure would hypothetically make, only an approximation for the purposes of ranking individuals and groups. However, it is still necessary to make them commensurable in a manner which is not arbitrary. This could be done if one treats the action of going from one square to another as an independent action. One could then calculate the number of possible moves from each square to give a number of movement actions. For the cardinal measure, this could then be added to the total freedom of action to give a new overall freedom total. For the social measure, it is more complicated; but one solution could be to create a movement value for each square by dividing the number of possible free moves by this same total in addition to the number of physically possible moves. This would create a figure for each square and these figures could then be added to the freedom of action for each square to give an overall score.

I have attempted, first, to elucidate the current problem of application and, second, to point out the deficiencies in previous attempts to rectify it. I have argued that these solutions fail because they leave a lot of room for contingent, unjustified judgments and that there is, therefore, a need to find a solution that minimises this. The solution I proposed tried to address this defect. In contrast to the other methods of application, this seems to provide a measure grounded in facts which are not contingent and which would lead to similar

measures regardless of who applied it. Whether it can be applied satisfactorily or not remains an open empirical issue. However, I see no obvious reason why this cannot provide an alternative solution to the problem of application – a solution that reduces the role of arbitrary judgement far more than existing theories.

References

- [1] Berlin, I. (1969) ‘Two Concepts of Liberty’, in I. Berlin, *Four Essays on Liberty*. Oxford: Oxford University Press.
- [2] Carter, I. (1999) *A Measure of Freedom*. Oxford: Oxford University Press.
- [3] Graeff, P. ‘Measuring Individual Freedom: Actions and Rights as Indicators of Individual Liberty’, in F. McMahon (ed.), *Towards a World Wide Index of Human Freedom*. Fraser Institute.
- [4] Kramer, M.H. (2003) *The Quality of Freedom*. Oxford: Oxford University Press.
- [5] Oppenheim, F.E. (2004) ‘Social freedom: Definition, measurability, valuation’, in *Social Choice and Welfare* 22: 175–185.
- [6] Pattanaik, P.K. and Xu, Y. (1990) ‘On Ranking Opportunity Sets in Terms of Freedom of Choice’, in *Recherches Économiques de Louvain* 56: 383–390.
- [7] Rosenbaum, E.F. (2000) ‘On measuring freedom’, in *Journal of Theoretical Politics* 12: 205–227.
- [8] Steiner, H. (1983) ‘How Free: Computing Personal Liberty’, in *Royal Institute of Philosophy Lecture Series* 15: 73-89.
- [9] Suppes, P. (1996) ‘The nature and measurement of freedom’, in *Social Choice and Welfare* 13: 183–200.

Interview: Mary Margaret McCabe

King's College London

What led you to choose Philosophy as a career?

Perhaps it was just that I couldn't stop – I started out as a classicist. When I was at school I studied Greek with the daughter of the wonderful Aristotle scholar and philosopher David Ross. We read two things together – Aeschylus' Agamemnon and Plato's Symposium. She showed me something about reading Plato that I never forgot: that he writes to be read with huge care, thought and reflection, and not just to tell us what to believe. When I was an undergraduate at Cambridge I had a supervisor, Denis O'Brien, who had the same view – that reading Plato just is doing philosophy, thinking about how thinking works; and then thinking about it all over again. He taught me how to see a page of Plato as though I had never seen it before, and to think about what might be going on there – so he showed me how these ancient texts could be absolutely fresh. Then I discovered that what completely captivated me about those texts were the arguments, the puzzles, the questions they were asking. The same, I found as I started teaching, was true of all sorts of parts of ancient philosophy – and I discovered at the same time just how much I loved the business of talking to others about the same puzzles and questions and texts – so perhaps it was teaching that really got me going: anything that I get right, I owe to those kinds of conversations with students and colleagues. The questions and the puzzles preoccupied me in ways that scholarship (I am afraid) never quite did... so I moved away from a Classics department in 1990 to KCL Philosophy Department – and that was, at that time and still, the most wonderful place to do philosophy, without the kind of 'philosophy as a competition' approach which I find so arid.

What do you think Philosophy is? What is Philosophy's role in the world?

So I think philosophy is just thinking about thinking, whatever form that seems to take, and however it is formalised. It is critical, highly demanding, precise, argumentative in the best way – and at its very best open-minded. This goes along, I think, with developing a kind of imagination that helps us think about how we are in the world, how we think about ourselves and others. All of this, I think, makes philosophy something that we do, not suffer; it makes it an activity, rather than a set of fixed views or beliefs – so I am more interested in philosophizing than in philosophies. If philosophy has a 'role in the world', it is to use that kind of critical care to make sure we think as best we can, and to make sure that our view of the world is rich and thoughtful. But I suppose

that brings responsibilities with it – failing to think clearly, especially in the public domain, is perhaps a kind of ethical lapse, rather like walking by the injured person on the other side of the street. I hope that doesn't sound just too horribly high-minded – but I am not convinced that there are firm lines of demarcation between the logical and the moral that makes it OK to see that an argument put forward in the public domain is horribly unsound – and not say so. I don't think, incidentally, that this means being aggressive or raging against others: indeed, I think that rage and aggression themselves damage how we think – and that too sounds horribly high-minded, doesn't it... Perhaps the best way to put it is that being civil to each other is an important aspect of how we do the subject.

What can we learn from Ancient Philosophy? What do you think ancient philosophy can teach us about ourselves as people?

One of the peculiar and difficult things about studying ancient philosophy is trying to get behind and away from the dominant assumptions we might bring to the interpretation of philosophy in the modern era. The ancients, at least in the Classical period, are not dominated, for example, by the problem of doubt – their accounts of knowledge are not, I think, driven by the question 'how can p be true?' but by the question 'how can p be false?'; equally not by the question 'how can we know p is true?', but rather by questions about explanation ('why did p come about?' etc.). Or – in ethics the dominant ancient worry is about the account we might give for virtue and goodness, rather than duty; and thinking about these thinkers gives us, I think, some kind of perspective on ourselves that closer thinkers don't allow (this is where the recent interest in 'virtue' ethics and epistemology comes from). At the same time, of course, they are somehow accessible to us all the same – the challenge is always to figure out how we negotiate between the strangeness of ancient philosophy and its familiarity – that negotiation, it seems to me, is something we can just keep right on doing... And I completely love reading a text with others, sitting and working out what on earth is going on, together. Perhaps that is made more tractable when one is dealing with a text in a very difficult and distant language?

Who do you prefer: Plato or Aristotle? Why?

Plato – it seems to me that the utter brilliance of how he writes, and how reading what he writes engages us in philosophy, has no equal. But in my old age I am also coming to appreciate Aristotle much more than I did – as a student I found him just hugely daunting; and I found the received opinion at that time, that Aristotle writes what he does as a set of rather crabbed lecture notes, rather dry. I have since realized – better late than never – that Aristotle is

reading Plato too, in the rich ways that he should be read, and that Aristotle's detail can be as hair-raisingly exciting as Plato's. Jonathan Barnes says that he prefers Aristotle to Plato because Plato's arguments are such rubbish, and Aristotle's so good – I think the relation between them is far richer than this, and that thinking about it as a relation illuminates what otherwise seems to be just a contrast between two completely different ways of tackling some of the same questions.

What challenges have you faced as a woman in Academia? Do you think Academia is becoming more equal?

This is tricky. When I was an undergraduate in the '60s I thought we had all of this sorted out, that there was a very short distance to go before we would have something that felt like equality – or, rather, something that felt as if gender differences were just insignificant when it came to academic matters. I have since – and perhaps especially recently, come to be rather more dispirited about how much progress we have made. There is a fair amount of formal equality; but the figures for the progress of women, especially in some STEM subjects and in Philosophy, are dismal. Jenny Saul has done some excellent work on how this may be partly a matter of implicit bias; and there is a great deal being discussed in philosophy now about how to improve things – so perhaps it is not dismal after all. And perhaps some of the things that bother me may be just individual and not gendered at all. Perhaps my readers should decide. I suffer badly from 'Imposter Syndrome' – from expecting that just today I shall be found out to be a complete imposter and not belong anywhere in philosophy at all. I also find fierce argument rather disturbing; even although I can probably be rather fierce from time to time when I get excited... I find the 'philosophy competition' just makes me anxious, makes me think again and again about whether what I have to say is worth saying and then end up saying nothing. (This was at its worst when I as a graduate student; in King's again, I found that a very kindly environment helps this to stop). But as a consequence, I do try very very hard to make sure that seminars I run and lectures I give are not intimidating; that people see that I can be just as wrong and mistaken as anyone, and most of all that thinking about these questions is exciting and fun and often funny, not some way of figuring out who is the cleverest (whoever it is, it won't be me). I read a recent interview with a philosopher in which he was offering a checklist of which philosophers were as 'smart as him' and which might be smarter (there was only one candidate): I find that sort of talk inhibiting and damaging. It should be possible to think that we are all in this together, should it not?

What do you think are the challenges the UK university system faces? Do you think they're surmountable?

I am not quite sure where to begin with this! The current funding system is not only fiscally unsustainable, but also deeply damaging, because it promotes a market ideology which is inappropriate to account for the value of the university system as a public good. The ways that this has worked out recently has, it seems to me, also seriously damaged how teaching in universities is done – even although hard-pressed departments try very hard to temper the bad effects of increased numbers of students. There are two other lurking problems (which are connected, of course, to the market ideology). The first is that we still haven't sorted out how to make access to university genuinely open; the problem of inequality is urgent, and I don't think we have worked out yet how to solve it (one thought may be to pluralise the sector more effectively: Martin McQuillan calls this the 'rainbow sector' and I think that is spot on). The second is that university governance has increasingly been taken away from academics and into the control of administrators; and I think this means that the issue of academic principle are not being voiced or heard.

What advice do you have for students who are currently studying Philosophy?

Enjoy it, and make sure not to turn it into a competition – do philosophy together if you can as a conversation, not a contest. Remember that it is a slow and laborious process, much slower than you might expect – don't be surprised if it takes you a week or more to read and understand an important article, don't think that there is somehow a right speed to do things. When you read, let the ideas settle and coalesce – give yourself time for thinking and rethinking. (That means, among other things, not doing philosophy at the last minute). And don't worry about it, don't worry if you don't get the point of something immediately, just allow yourself time to think things through. Just take it easy!

MMM 2.11.14

Issue 8(1)
Autumn Conference 2014

Subsistence as a Basic Human Right:

A defence of Henry Shue

Sean Butler

A Set of Objections to

David Benatar's Anti-Natalism

Franco Palazzi

*How Convincing is the
Levelling Down Objections
to Egalitarianism?*

Joel Griffith

Killing the cat with too much cream?

A review of Donaldson and Kymlicka's

Zoopolis

Collins Tahzib

*Kurt Gödel's Proof of the Ideality of
Time*

Lukas Clark-Member

*Michael Smith's 'Moral Problem' and
a Defence of Metaethical Externalism*

Laura Nicoara

*On the Moral Status of Objectification
in Pornography*

Carl-Otto Frietsch

Can We Practically Measure Freedom?

Alexander Toucher

Interview

Mary Margaret McCabe



Journal of the British Undergraduate Philosophy Society

(Print) ISSN 1748-9393
(Online) ISSN 2051-5359