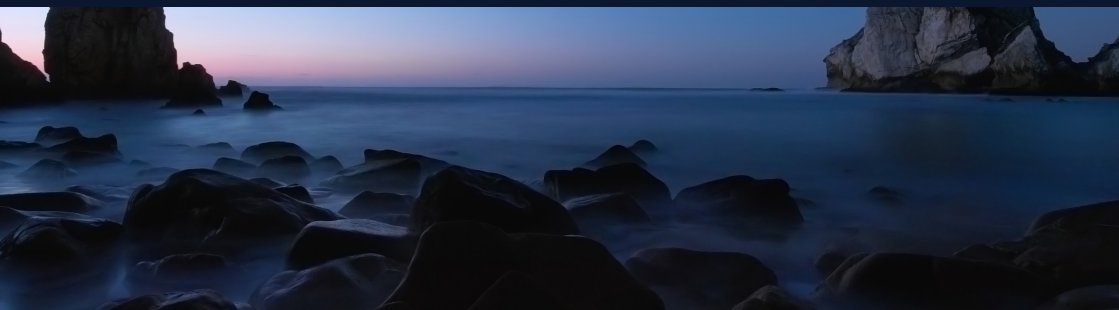


Issue 7(1) Autumn Conference 2013

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

*British
Journal of
Undergraduate
Philosophy*



Editor-in-chief: Dino Jakušić
University of Warwick

Journal of the British Undergraduate Philosophy Society



Greetings!

Your undergraduate studies might soon be at an end, and perhaps you are wondering about your next step. I am writing to let you know about the one-year MSc programmes that we offer in the Department of Philosophy, Logic and Scientific Method here at the LSE:

MSc Philosophy of Science
MSc Economics and Philosophy
MSc Philosophy and Public Policy
MSc Philosophy of the Social Sciences

We offer a type of philosophy that engages with and aims to make a difference to the practice of the sciences and policy-making. We have designed a curriculum for each MSc that is thoroughly interdisciplinary and that teaches you how to combine philosophical reflection with scientific insights and policy applications. Discussions are conducted in small seminars with students from a variety of backgrounds and from all over the world.

Our teaching is research-led: courses cover current research in the field and are taught by world-leading experts in the field. But don't just take our word for it. In the Philosophical Gourmet Report 2011 the Department was ranked 1st in the world for philosophy of the social sciences, joint 2nd in the world for philosophy of science, joint 4th in the world for decision, rational choice and game theory, and joint 5th in philosophy of physics. In addition we are placed well in applied ethics and political philosophy and have recently made some excellent new appointments in these areas. The Department scored first in the Guardian University Guide 2014 for Philosophy.

Upon graduation our students pursue a variety of careers. For those who wish to pursue an academic career our MScs provide an ideal springboard for a PhD. But an MSc also opens many other doors. Our students go on to work as consultants, journalists, civil servants, teachers, investment bankers, or researchers in NGOs or international organisations.

If you wish to hear more about these opportunities, do visit us on the LSE website at www.lse.ac.uk/philosophy. We would also be very happy to talk to you in person. Feel free to drop us a line at philosophy-dept@lse.ac.uk – make sure to mention the programme or programmes that you are interested in, and one of our coordinators will be in touch with you shortly.

I look forward to hearing from you!

Yours sincerely,

A handwritten signature in black ink that reads 'Luc Bovens'.

Professor Luc Bovens
Head of Department

British Journal of Undergraduate Philosophy

www.bups.org/bjup-online

All rights reserved.

© 2013 British Undergraduate Philosophy Society, and the authors.

Cover image by Wikimedia Commons user 'Rnbc' used under the terms of the GNU Free Documentation License. Photograph of Jerry Fodor by Wikipedia user 'Pealco' released into public domain. Photographs of Huw Price and Shalom Lappin used with permission.

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

British Journal of Undergraduate Philosophy

The Journal of the British Undergraduate Philosophy Society

Issue 7(1) Autumn 2013

Editor-in-chief:
Assistant editor:

Dino Jakušić
Daniel Bregman

Manuscript editors:

Lukas Apse
Christopher Wing Hei Cheung
Luke Deveraux
Liberty Fitz-Claridge
Michael Lyons
Sophie Osiecki
Dan Treger
James Wood

Commissioning editors:

Megan McCarthy
Julie Lee

www.bups.org/bjup-online

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

Editorial

With the Autumn 2013 Conference approaching it is time for *BJUP* 7(1) volume to appear as well. It is interesting to observe trends in submissions and what seems quite popular this year is logic and a discussion of paradoxes. I guess it is a nice note on which to end a (quasi-)paradoxical situation in which British Undergraduate Philosophy Society and its Journal are run by philosophy postgraduates (Michael and myself). While this issue, and the conference which it accompanies, will without a doubt be interesting to all I think it is fair to say that they will remain in special memory for Michael and myself since they will probably be our last. So without further ado I would like to thank everyone who has (successfully or unsuccessfully) sent their work to the forthcoming conference and/or the journal issue and everyone who has helped to make this issue possible from the technical side of the story.

First of all, I would like to especially thank to the Assistant Editor, Daniel Bregman, both for his excellent work in modernising the typesetting and compiling of the Journal and also on the tremendous amount of work he put into producing this issue. Anyone who has been familiar without our previous \LaTeX practices will agree that the \LaTeX side of the journal no longer feels like it is kept together with string and some glue.

Second of all, I would like to thank the whole of the manuscript team for the excellent and swift job they have done in typesetting, reference-checking and proof-reading the journal, especially since some of them were completely new to \LaTeX and *BJUP* (also special thanks to Dan Treger for participating in the horror story which is the selection of papers for publication).

It would also be wrong not to thank our commissioning editors for the tremendous job they have done with receiving, anonymizing, de-anonymizing and sending papers to external reviewers and responding to sudden requests for the most obscure information which Daniel, Michael or me would remember we urgently needed at a certain point.

This leads me to thanking our external reviewers: Antonia Rebecca Simpson, Miranda del Corral, Farah Abdessamad, Dr. Bettina Bohle, Maria Avxentevskaya, Sarah-Louise Johnson, Nathan Oseroff, Jonathan Banks, Nahuel Snajderhaus, Dr Anthony Everett, Alizera Sayadmansour, Takashi Oki, Dr Neil Turnbull, Sumeyye Parilder, Željko Mančić, Laura Pelegrin, Janelle Potzsch, Josh Milburn, Dr Ian Church, Jason Turner, Biswanath Swain, Dr James Stark, Dr Angelo Cei, Katherine Hawley, Dr Nigel Pleasants, Jonathan Head, Robert Pezet, Mats Volberg. Thank you all for volunteering to review our submissions

and special thanks to all of you which have been working with us for several years now.

Finally, I most certainly have to thank our president, Michael Lyons, for leading us through the previous, quite difficult years, and for putting tremendous amount of work both in the Society in general and in the journal by being, as he described himself, “the journal handyman”.

I hope to see as many of you as possible at the conference or later in life, possibly in the academia.

DJ

Contents

- i. A message from the LSE
- v. Editorial
- 1. **Properties, Predicates, and Evans' Argument against Vague Objects**
Chris Blake-Turner
- 9. **The Causal Theory of Reference, Sameness-Relations and Quasi-Constructivist Essentialism**
Sam Carter
- 19. **Verificationism and Relativism: A New Defence of Metaphysics**
Atticus Gatsby
- 26. **A Critique of Jeffrie Murphy on Resentment**
Bjorn Wastvedt
- 39. **Tackling the Repugnant Conclusion**
Helen Weir
- 51. **Socrates' Philosophical Mission in the *Apology*: Irony and Obedience**
Corrado Musumeci

67. **Environmental Philosophy and the Teleological Argument: A Polemic**
Thomas E. Randall
80. **On the Relation Between the Form of Beauty and *Eudaimonia* in Plato's *Symposium***
Carl-Otto Frietsch
90. **Hegel's *Philosophy of Right*: Civil Society and International War**
Joel Mason
98. **Are Scientific Revolutions Irrational?**
Sophie Osiecki
104. **Interview: Jerry Fodor**
111. **Interview: Huw Price**
120. **Interview: Shalom Lappin**

Properties, Predicates, and Evans’ Argument against Vague Objects

Chris Blake-Turner

Durham University

Properties are ways objects are.¹ An electron’s being negatively charged is a matter of its being a certain way, and this is a matter of its possessing the property of $-1e$ charge. Predicates, on the other hand, are linguistic units, which we ascribe to entities.² They can apply or fail to apply depending on the nature of the entities to which they are ascribed.³ Accordingly, when we say of a certain electron that it ‘is negatively charged’ the predicate truly applies in virtue of the electron’s possessing the relevant property.⁴ A question of fundamental importance is: how are properties and predicates related?⁵ It might be initially attractive to suppose that they are coextensive. I will first claim that this is not the case. After rejecting a weaker correspondence principle, I will then argue that the crucial predicate in Evans’ argument against vague objects can be truly applied without positing a corresponding property. For this reason, the argument fails in its original form. This is a good example of the danger of taking the relation between properties and predicates for granted; we should avoid this and any other practice that “derives ontology from linguistic prejudice”.⁶

¹I intend this characterisation to be neutral with the respect to the trope-universal debate. For simplicity, I will talk in the essay *as if* properties are universals, but the same points can be made for tropes by making the relevant changes.

²Henceforth the scope of the essay is limited to *meaningful* predicates.

³A realist framework is being assumed.

⁴I am not here concerned with what exactly *applying in virtue of* consists in. Furthermore, I am assuming a weak truthmaker principle, such that when a predicate truly applies, there must be *something* about reality in virtue of which this is the case, though this need not be a property. If it is preferred, such talk of properties can be construed as short for properties *qua* aspects of entities, or properties *qua* constituents of states of affairs. See further: Heil [5] pp. 61ff.; Williamson [15] pp. 705–706; Armstrong [2] pp. 89ff.

⁵This question is ambiguous: it might mean, what is the specific nature of the relation that holds between predicates and properties? Or it might mean, between which properties and predicates does that relation hold? This essay deals with the latter. Cf. footnote 4.

⁶Williamson [15] p. 706.

1 Coextension

When the predicate ‘is woolly’ truly applies to a jumper, this is because the jumper really is woolly. Perhaps this close connection between properties and predicates can be explained by their being coextensive, such that for every predicate there is a property and *vice versa*:

- (1) P^* is a predicate if, and only if, there is a corresponding property P .

However, it is easy to see that (1) is false, for there are predicates which have no corresponding properties, no “ontological correlates”.⁷ For instance, a version of Russell’s paradox arises from considering the predicate ‘is non-self-exemplifying’: does the property it purports to pick out exemplify itself?⁸ If it does, then it cannot be *non*-self-exemplifying, and yet that means it does *not* exemplify itself; if it does not, then it is non-self-exemplifying, which means it *does* exemplify itself. Since there is at least one predicate to which no property could correspond, on pain of contradiction, (1) is false.⁹

2 A weaker correspondence principle

Clearly predicates and properties are not coextensive, but maybe the link is still very close, as captured by something like the following weaker correspondence principle:

- (2) If a predicate, P^* , truly applies to an entity, then (i) the entity possesses some property (or relation), P , in virtue of which P^* truly applies, and (ii) *every* entity to which P^* truly applies possesses P .¹⁰

I do not wish to take issue with (2.i) for now,¹¹ but (2.ii) needs immediate scrutiny. There is something initially plausible about it, and it seems to be true

⁷Williamson [15] p. 699.

⁸Lowe [10] p. 364.

⁹(1) is also false when read right-to-left, as there are plausibly properties with no corresponding predicates, no ‘linguistic correlates’. See further Armstrong [1] p. 171.

¹⁰Cf. Heil [5] p. 26.

¹¹Since ‘is non-self-exemplifying’ never truly applies – all entities are examples of themselves – it does not threaten (2.i).

for some predicates: ‘is woolly’ truly applies both to my jumper and my hat because they both have the same property, namely *woolliness*. However, there are many predicates for which (2.ii) is false. Consider ‘is a projectile’, which can be truly predicated of bullets, arrows, baseballs and human cannonballs. The predication is true of all these objects in virtue of some property that each possesses, but it is very unlikely that each relevant property is *identical*. This is because objects can be projectiles *in different ways*, such as by being fired from a gun, shot from a bow, thrown by a pitcher, or ejected from a cannon, and the relevant property will differ accordingly. Predicates allow users to perform the useful linguistic task of abstracting away from individual properties and grouping together relevantly similar ones, but this cognitive abstraction does not require the existence of ‘higher-level’ properties in the world.¹²

3 Evans’ argument against vague objects

The above discussion is important not because many philosophers explicitly endorse (2) – they do not – but because not considering carefully whether a predicate has the ontological correlate it seems to can lead to bad argumentation. For example, consider Evans’ famous argument against vague objects.¹³ Let a and b be precise designators, and suppose for *reductio* that it is indeterminate whether a is identical with b .¹⁴ Using ∇ as an indeterminacy operator, we have:

$$(3) \quad \nabla(a = b)$$

From (3) we can ascribe to b the property $[\nabla(x = a)]$, or *being such that it is indeterminate whether it is identical with a* :

$$(4) \quad [\nabla(x = a)]b$$

But it is very plausible that whether a is identical with a is *not* indeterminate:¹⁵

$$(5) \quad \neg\nabla(a = a)$$

¹²Heil & Robb [6] pp. 179–183. Cf. Wittgenstein’s [16] discussion of what games (do not) have in common, §§66ff.

¹³Evans [3] p. 208.

¹⁴A *precise designator* denotes the same entity whenever it denotes anything at all. See further: Noonan [11] p. 4; Lewis [7] pp. 128–129; Tye [14] pp. 556–557.

¹⁵See further Lowe [9] p. 70, who suggests that “determinate *self*-identity is a necessary condition of entityhood”.

Thus:

$$(6) \quad \neg[\nabla(x = a)]a$$

From Leibniz's law and the fact that b has a property that a lacks, namely *being such that it is indeterminate whether it is identical with a* , we can derive the following:

$$(7) \quad \neg(a = b)$$

(7) does not directly contradict (3), but it can be made to: given that we assumed it was *definitely* indeterminate whether a is identical with b , the conclusion that a is not identical with b is also definite.¹⁶ So, using Δ for a determinacy operator:

$$(8) \quad \Delta\neg(a = b)$$

Which yields, by the duality of Δ and ∇ :¹⁷

$$(9) \quad \neg\nabla(a = b)$$

(9) *does* straightforwardly contradict (3) and so the argument seems to show that identity statements with precise designators cannot be indeterminate on pain of contradiction. In other words, vagueness has to be semantic rather than ontic.

While Evans' argument has been questioned in many ways,¹⁸ I think it is invalid because it assumes an enthymematic premise that is false: to get from (3) to (4) it assumes that if we can predicate 'is such that it is indeterminate whether it is identical a ' then there must be a corresponding *property* $[\nabla(x = a)]$.¹⁹ That assumption is false because (2.i) is false. Not only are there some predicates, like 'is a projectile', which can truly apply to different entities in virtue of different ontological correlates, but there are also some predicates that can truly apply despite having *no* property as an ontological correlate.²⁰ To see this, let us consider predications of identity.

¹⁶Williamson [15] p. 707.

¹⁷Noonan [12] p. 17.

¹⁸For example: Garrett [4] pp. 131–133; Lowe [9] pp. 63–67; Parsons & Woodruff [13].

¹⁹Cf. Lowe [9] p. 65.

²⁰Possible candidates for the ontological correlates might be the entities themselves, or a state of affairs. See further: Heil [5] pp. 61–74; Armstrong [2] pp. 88–94.

There is a minor complication to be taken into account at this point: predications of identity are taken to signify a *relation*, rather than a property, and in general predicates seem to be able to have relations as their ontological correlates. So ‘is taller than’ might have the relation of *being taller than* as its ontological correlate. If properties are ways things are, then relations are ways things are *with respect to each other*. While in general relations can be taken with ontological seriousness, I want to deny this in the case of identity. When ‘*a* is identical with *b*’ is true, the ontological correlate is not a substantive relation of identity *in the world*; identity is a purely formal relation that ‘comes for free’, ontologically speaking. Why deny the identity relation substantive status but allow it to *being taller than*? One very good reason is that if the former is substantive, then an infinite regress threatens: let *a* be identical with *b*; then there is a substantive identity relation, i_1 , holding between *a* and *b*; but by the reflexivity of identity, i_1 is identical with i_1 ; then there is a *further* substantive identity relation, i_2 , holding between i_1 and itself; but by the reflexivity of identity, i_2 , is identical with i_2 ; and so on...²¹ This regress “is either vicious or at least viciously uneconomical”.²²

Returning to Evans’ argument, recall the predicate ‘is such that it is indeterminate whether it is identical with *a*’. What is the ontological correlate of this predicate? The argument needs it to be a property to go through. It is important at this stage *not* to start discussing a separate relation of ‘indeterminate identity’. As Williamson points out, responding to the argument by reinterpreting ‘=’ as something other than identity is not really responding to it at all.²³ Keeping our focus on identity, it was argued above that whether *x* is identical with *a* “flow[s] from the natures” of the relata.²⁴ The same is plausibly true of ‘is such that it is indeterminate whether it is identical with *a*’, the truth or falsity of which need not depend on some *property* [$\nabla(x = a)$], but which stems from the very natures of whatever it is predicated of and *a*. To put it another way, suppose there *were* a property [$\nabla(x = a)$]. Suppose further that it is to be decided whether *b* is such that it is indeterminate whether it is identical with *a*. Suppose that every relevant fact about *b* and

²¹Cf. Armstrong [2] pp. 109. What I have been calling ‘formal’ relations, he calls “internal relations”, which are “nothing over and above their terms”.

²²Ibid. p. 108.

²³Williamson [15] p. 709. Besides, indeterminate identity is *not* subject to the same regress of identity, since the former is presumably identical, rather than indeterminately *identical*, with itself; this causes no problems since we are taking identity to be formal.

²⁴Armstrong [2] p. 57.

a are known, *except* whether b possesses the property $[\nabla(x = a)]$. It seems very plausible that in this situation it would *already* be possible to tell whether b is such that it is indeterminate whether it is identical with a or not.²⁵ In that case, b 's possessing $[\nabla(x = a)]$ would be completely irrelevant. This does not mean that we *cannot* posit the property's existence, but why bother if it is redundant, "of a highly dubious nature and completely without empirical significance"?²⁶ Given this, the ontological correlate of the predicate 'is such that it is indeterminate whether it is identical with a ' is not likely to be a property (or relation).

What is the upshot of this? Evans' argument as it has been set out above can be resisted: we can agree with (6) that a does not possess the property $[\nabla(x = a)]$. However, since there is no reason to believe that there *is* such a property, b also does not possess it, and so (4) can be rejected.²⁷ This leaves Leibniz's Law and the possibility of vague objects intact. This does not mean the overall strategy of argument need be rejected, however. Drawing on Shoemaker's Brown/Brownson thought experiment, Noonan has constructed a similar argument where the crucial predicate is 'is such that it is indeterminate whether he is thin after the transplant',²⁸ which certainly seems to have a better claim to having a property as an ontological correlate than the predicate involving identity. There are ways to try and respond to this reconstructed argument,²⁹ but it is not my purpose to go into them here. Instead I want to urge that if indeed Evans' original predicate does not have a property as an ontological correlate, then future efforts should rather be focused on the argument in Noonan's form, which does not seem illegitimately to ascribe a property merely on the basis of a predication.

²⁵I take it that this fact is not question-beggingly included amongst all the relevant facts, but rather *stems from* them. Furthermore, the Identity of Indiscernibles need not be assumed, since the relevant facts will presumably include spatiotemporal ones.

²⁶Lowe [9] p. 65.

²⁷Ibid.

²⁸Noonan [12] p. 17. Brown's brain is transplanted out of his fat body into Robinson's thin one. Call the person after the transplant (Brown's brain and Robinson's body) Brownson. Plausibly, it is indeterminate whether *Brown* is thin after the transplant, but it is determinate whether Brownson is (he is).

²⁹For example: Lowe [8]; Lowe [9] pp. 67–69.

4 Conclusion

I have argued that the relation between properties and predicates is complicated. One thing seems certain: they are not coextensive, since ‘is non-self-exemplifying’ cannot correspond to a property on pain of contradiction. Perhaps instead a weaker correspondence principle such as (2) describes the relation between them. This seems to be true for some predicates such as ‘is woolly’, but (2.ii) seems false in the case of others, like ‘is a projectile’. Things of which ‘is a projectile’ truly applies do have *a* property in virtue of which it applies, but it is not plausible to suppose it is the *same* property for *every* true predication; the properties may well be similar in some regard, so that it is convenient for us to group them under a common predicate.

It is important to consider whether a predicate deployed in an argument really has the ontological correlate it purports to. In looking at Evans’ argument against vague objects, I argued that the predicate ‘is identical with *a*’ can truly apply even when there is no corresponding substantive relation. This is because identity is a formal relation, which stems from the nature of its relata: whether *a* is identical with *b* is due to the natures of *a* and *b* themselves and not some further, substantive entity in the world. It seems likely that the crucial predicate of the argument, ‘is such that it is indeterminate whether it is identical with *a*’ also lacks a property or relation as an ontological correlate. For this reason, Evans’ argument illegitimately derives a substantive property from a true predication and so fails. However, a very similar argument can be made without using a predicate involving identity, and it is *this* argument that merits the real philosophical attention. Timothy Williamson has suggested that “it is not obvious why the general issues [to do with vague objects] should be thought to turn on the very special relation of identity”. If I am right, they do not.

More generally, Evans’ argument warns us to scrutinise the predicates we use in philosophical discourse; if they do not have the ontological correlates they purport to, the arguments we deploy them in might not work as we intend them to.

References

- [1] Armstrong, D.M. (1972) ‘Materialism, Properties and Predicates’, in *Monist*, 56(2): 163–76.

- [2] Armstrong, D.M. (1989) *Universals: An Opinionated Introduction*. Boulder, CO: Westview Press.
- [3] Evans, G. (1978) 'Can There Be Vague Objects?', in *Analysis*, 38(4): 208.
- [4] Garrett, B.J. (1988) 'Vagueness and Identity', in *Analysis*, 48(3): 130–34.
- [5] Heil, J. (2003) *From an Ontological Point of View*. Oxford: Clarendon Press.
- [6] Heil, J. & Robb, D. (2003) 'Mental Properties', in *American Philosophical Quarterly*, 40(3): 175–96.
- [7] Lewis, D. (1988) 'Vague Identity: Evans Misunderstood', in *Analysis*, 48(3): 128–30.
- [8] Lowe, E.J. (1997) 'Reply to Noonan on Vague Identity', in *Analysis*, 57(1): 88–91.
- [9] Lowe, E.J. (1998) *The Possibility of Metaphysics: Substance, Identity, and Time*. Oxford: Clarendon Press.
- [10] Lowe, E.J. (2002) *A Survey of Metaphysics*. Oxford: Oxford University Press.
- [11] Noonan, H.W. (1982) 'Vague Objects', in *Analysis*, 42(1): 3–6.
- [12] Noonan, H.W. (1995) 'E.J. Lowe on Vague Identity and Quantum Indeterminacy', in *Analysis*, 55(1): 14–19.
- [13] Parsons, T. & Woodruff, P. (1995) 'Worldly Indeterminacy of Identity', in *Proceedings of the Aristotelian Society*, New Series, 95(1): 171–91.
- [14] Tye, M. (1990) 'Vague Objects', in *Mind*, 99(396): 535–57.
- [15] Williamson, T. (2003) 'Vagueness in Reality', in Loux, M.J. & Zimmerman, D.W. (eds.) *The Oxford Handbook of Metaphysics*. New York, NY: Oxford University Press.
- [16] Wittgenstein, L. (2009). *Philosophical Investigations*. 4th ed. G.E.M Anscombe, P.M.S. Hacker & J. Schulte (trans.). P.M.S. Hacker & J. Schulte (eds.). Chichester: Wiley-Blackwell.

The Causal Theory of Reference, Sameness-Relations and Quasi-Constructivist Essentialism

Sam Carter

University of Edinburgh

Introduction

This paper will: (1) introduce Putnam’s causal theory of reference (hereafter, ‘CToR’), and show how essentialism concerning natural kinds is derived from it; (2) argue that there is no non-arbitrary, privileged ‘sameness’-relation serving the role required by CToR; (3) introduce five desiderata for an account of natural kinds; and (4) reconstruct a version of CToR which accounts for the criticisms in §2 and satisfies the desiderata of §3.

1 Putnam’s Causal Theory of Reference

Putnam’s causal theory of reference is primarily intended to provide an alternative to the view that extension is determined (solely) by the psychological states of speakers. Nevertheless, it has been widely argued¹ that an essentialist account of natural kinds can also be derived from CToR. According to CToR, the extension of a natural kind term is determined by a triadic, cross-world ‘sameness’-relation as follows.

(CToR) For a relevant kind-term ‘ K ’, sample ‘ a ’, property-type ‘ ϕ ’ and object ‘ x ’, x falls in the extension of K iff x bears the same-with-respect-to- ϕ relation to a .

It follows from CToR that intensions, *qua* psychological states of speakers, do not fully determine extension; semantic content is (partially) external. Two speakers can be psychologically identical regarding a term K , yet each may have samples for K which do not bear the relevant sameness-relation to the

¹Cf. Ellis [5]; Soames [25].

other's sample. Consequently, their extension of K would differ. This phenomenon is exemplified in Putnam's Twin Earth thought experiment.²

Given the above, the derivation of essentialism from CToR is relatively simple. Let essentialism be the view that F is an essential property of kind K iff membership of K necessarily entails possession of F :

(Essentialism) F is essential to K iff $\forall x(\Box(x \in K \rightarrow Fx))$.

This view should be distinguished from the claim $\forall x(x \in K \rightarrow \Box Fx)$, which asserts that membership of K entails necessary possession of F . The latter is a claim about individual essences of members of K , rather than essential properties of membership of K . As noted, sameness is a cross-world relation; that is, it can hold between objects in different possible worlds.³ Thus the conjunction of sample a in the actual world with the relation of sameness is sufficient to determine the extension of K in all possible worlds. Consequently, K will rigidly designate; it will refer to the same things (with-respect-to- ϕ) in all possible worlds in which it refers.⁴ From this we can derive essentialism. In all possible worlds, if an object x with property ϕ is a member of K in virtue of bearing the same-with-respect-to- ϕ relation to the sample a , then x cannot lose ϕ without ceasing to be a member of kind K . Equivalently we can say, 'necessarily, all members of K possess ϕ ', or ' ϕ is an essential property of K -ness'.

2 The Relevant Sameness-Relation Problem

CToR involves a triadic sameness-relation which holds between two objects and a particular property. Accordingly, no two objects can be said to be the same *simpliciter*. They may be the same-with-respect-to- ϕ , while differing-with-respect-to- ψ . §2 will argue that CToR possesses no non-*ad-hoc*, non-question-begging means of determining the relevant set of sameness-relations which must hold between a sample a and an object x for kind membership. Where 'same_L' denotes sameness-with-respect-to-liquid, and 'this' indexically picks out a sample, Putnam expresses CToR as the position:

²Putnam [20] pp. 223–25.

³Ibid. p. 323.

⁴Ibid. p. 231; Kripke [14].

(For every world W)(For every x in W)(x is water $\equiv x$ bears same_L to the entity referred to as ‘this’ in the actual world).⁵

Since any sample of water possesses numerous properties *qua* liquid, a sample can stand in multiple, distinct sameness-relations (same-with-respect-to-colour; -viscosity; -thirst-quenching-ness; etc.). Some means of distinguishing ‘relevant’ sameness-relations must be identified. Putnam explicitly denies that macro-level properties such as colour or viscosity can determine extension. They are elements of the stereotype alone, which serves to identify a sample, and thus provides necessary, yet insufficient conditions for being water. Rather, relevant sameness is determined by a privileged subset of sameness-relations: sameness-with-respect-to-microstructure.⁶ In general, for any kind K with sample a , x is a member of K iff x bears the sameness-with-respect-to-microstructure relation to a . Yet the same problem recurs. As with liquids, objects possess many different micro-structural properties, meaning they can be micro-structurally similar and dissimilar in many different respects (sameness-with-respect-to-elements; -arrangement; -reactivity; etc.). The metaphysical descent to micro-structure cannot distinguish a unique sameness relation.

An intuitive solution is to say that objects must be the same-with-respect-to all (intrinsic) micro-structural properties; that is, they must be micro-structurally qualitatively identical. However, this is both too strong and too weak a condition. Demonstrating the former, the water dimer ($(\text{H}_2\text{O})_2$) is a water cluster which differs from H_2O in that it is formed of two H_2O molecules joined by a weak hydrogen bond.⁷ $(\text{H}_2\text{O})_2$ is not qualitatively identical to H_2O , yet they belong to the same kind, water. Thus qualitative identity is too strong. Yet, qualitative identity is simultaneously too weak. H_2O molecules in their gaseous state are micro-structurally identical to H_2O molecules in their liquid state. Yet, if the stereotype of ‘water’ consists of macro-level sameness-of-liquid-relations, it is a necessary condition of water that it is liquid. Qualitative identity thus includes objects which are not of the kind, and excludes objects which are. It may be possible to identify an arbitrary set of sameness-relations which hold between all-and-only those objects which occur in ordinary language extension of the term (e.g. H_2O , its polymers and isotopes). Yet such an endeavour would be highly *ad hoc* in nature and contrary to the direction of explanation in CToR. It would first assume the extension of the term, then

⁵Putnam [20] p. 231.

⁶Ibid. p. 239.

⁷Sienko [23] pp. 344–45; Housecroft [11] pp. 210–11.

modify the theory of reference to fit it. Finally, Putnam could establish kind-extension by requiring that kind members must be alike-with-respect-to-all-*essential*-properties. Yet doing so would beg the question, since the sameness-relation was intended to determine the essential properties of a kind. Mellor⁸ and Salmon⁹ have similarly argued for the conclusion that CToR does not prove essentialism, but rather must take it as an assumption. Furthermore, even if essentialism is accepted as a premise rather than conclusion of CToR, Sterelny¹⁰ notes that any given sample will be a member of multiple natural kinds. It will therefore possess multiple essential properties which must still be distinguished.

3 Desiderata for an Account of Natural Kinds

Before proposing a resolution of the problem of determining relevant sameness-relations, we will identify five desiderata for a satisfactory account of natural kinds:

- (1) *Projectability*: A property is projectable if it supports inductive inferences. It is a standard necessary (though insufficient) condition of natural kinds that they should distinguish between those properties which are projectable, and gruesome Goodman-type properties, which are not.¹¹ If ϕ is a property determining membership of K , ϕ must be projectable (thereby allowing K to figure in successful inductive inferences).
- (2) *(Weak) realism*: Related to (1), natural kinds constitute classes which ‘carve nature at its joints’. Accordingly, under any satisfactory account, the boundaries of kinds ought to be determined by objective features of reality.¹² If ϕ is a property determining membership of K , ϕ must be a mind-independent, ‘natural’ property. This condition falls short of strong realism, since it does not suppose that kinds themselves constitute mind-independent objects, only that the properties determining their membership do.

⁸Mellor [17].

⁹Salmon [22].

¹⁰Sterelny [24].

¹¹Quine [21]; Boyd [1].

¹²Mill [18]; Quine [21]; Hacking [8] p. 110.

- (3) *Extensional change across (radical) theory change*: Kuhn¹³ criticises CToR on the grounds that it entails that extension is invariant through radical theory change. Intuitively, the extension of the term ‘planet’ changed after the Copernican revolution to include Earth. Previously, ‘planet’ had meant bodies orbiting the Earth; latterly, it meant bodies orbiting a star.¹⁴ Yet, under CToR, if ‘planet’ is a natural kind term at all, then, given a sample (e.g., Mars) and appropriate sameness-relations, it will include Earth both pre- and post-revolution. Kuhn’s issue is thus primarily with CToR’s semantic externalism, since it implies that a community’s theoretical beliefs can radically alter, while their terms’ meanings remain unchanged. A satisfactory account should allow for radical changes in theory to affect kind terms’ extensions.
- (4) *Scientific discovery*: Kuhn rejects CToR in favour of a nominalist view, whereby kinds’ extensions are fixed via description within a theoretical setting. The undesirable consequence of such descriptivist positions is that, instead of scientific investigation discovering the nature of kinds (as it does under CToR), they become *post hoc* classes constructed by human stipulation *after* a scientific breakthrough. We do not find out that water = H₂O, but rather stipulate it. Any satisfactory account should allow for discovery of kinds’ properties.
- (5) *Non-cross-cutting kinds*: Following Hacking¹⁵ and Ellis¹⁶, for a class to classify as a kind it must not ‘cross-cut’ any established kind. Two sets cross-cut if they are not disjoint sets and neither is a proper subset of the other (that is to say, they share some but not all members). This condition allows real kinds to be arranged in a hierarchical taxonomy.

4 Community Interests and the Causal Theory of Reference

§4 will propose an interest relative revision of CToR (hereafter, CToR_{INT-REL}) which will be demonstrated to resolve the sameness problem of §2, while simultaneously satisfying desiderata (1)-(5) and retaining CToR’s essentialism and externalism:

¹³Kuhn [15], [16].

¹⁴Ibid. pp. 312–13.

¹⁵Hacking [9].

¹⁶Ellis [5].

(CToR_{INT-REL}) For a community C , with interests i , let R^i be the set of sameness-relations relevant to i , i.e.

$$R^i = \{\text{same-with-respect-to-}\phi^1; \dots; \text{same-with-respect-to-}\phi^n\}$$

For any object ‘ x ’, kind term ‘ K ’, and sample ‘ a ’, x falls in the extension of K for C^i iff x bears R^i to a .

Let an interest i be a practical goal shared by a set of agents C . A given sameness-relation, ‘same-with-respect-to- ϕ ’, is ‘relevant’ relative to i iff the property ϕ is a factor in determining C ’s ability to achieve their interests. For example, for a community of astronomers, with the goal of predicting planetary movement, sameness-with-respect-to-stability-of-orbit will be relevant to determining membership of the kind ‘planets’. Conversely, sameness-with-respect-to-material-composition will not, since a planet’s substance does not have a significantly large impact upon predictions of its trajectory. Stable orbit is one of three official necessary conditions of planet-hood, while substance is unmentioned.¹⁷ Similarly, for a community of botanists, who share the interest of categorising plants within a phylogenetic-based taxonomic system, onions and lilies fall within a single kind, the family Liliaceae. Relative to a community of chefs, however, they do not, since sameness-with-respect-to-ancestor will be irrelevant to gastronomic interests. Instead, we might expect to find kinds determined by sameness-with-respect-to-taste. If desired, CToR_{INT-REL} can support a version of Putnam’s division of linguistic labour. By fixing the extension of certain terms for all speakers based on the interests of a sub-class of experts, CToR_{INT-REL} will be able to maintain the intuitive extensions of technical terms such as ‘positron’, ‘ganglion’, ‘elm’ and ‘beech’, even if the properties determining them are unimportant to most speakers’ goals.

Quasi-Constructivism. CToR_{INT-REL} is a quasi-constructivist view insofar as the membership (and even existence) of kinds can vary across different communities with different interests. This allows for satisfaction of desideratum (3), since extension can vary over time in accordance with Kuhnian changes in theory. When a scientific field undergoes radical change, the interests and aims of researchers are liable to change. For example, after Pasteur’s discovery of the role of germs in disease, the medical community adopted the aim of identifying and eliminating particular disease-causing micro-organisms. Accordingly, sameness-with-respect-to-bacterial-origin became relevant in classifying

¹⁷International Astronomical Union [12].

diseases. This led to changes in medical kinds reflected by $CToR_{INT-REL}$. For instance, a number of diseases previously thought distinct were collectively identified as tuberculosis. The continuity of broad domains of enquiry (such as medicine) across ‘paradigm shifts’ can be explained by sustained, broad-ranging goals which remain unchanged.

Though $CToR_{INT-REL}$ exhibits constructivist features, constructivism in its stronger forms, which hold kinds to be dependent upon speakers’ beliefs, is denied. There exist two strong reasons for treating interests as pragmatic goals rather than identifying them with agents’ mental states. Firstly, goals can be shared by large groups of people, while it may be that no two individuals are entirely alike with respect to psychological states. Analysing interests as practical goals allows extension to remain stable across the community, rather than varying from speaker to speaker. Secondly, it ensures that speakers’ beliefs play no role in kind membership. Kinds, though relative to communities, are not consciously or actively constructed by them. $CToR_{INT-REL}$ remains a form of semantic externalism, since the extension of kind terms is dependent on agent-independent factors; speakers may not be aware which kind-determining properties are relevant to their interests. This allows for the real discovery of facts about kinds natures, satisfying (4).

Experimental Realism. In making kind membership dependent upon the properties a community must manipulate to achieve success, $CToR_{INT-REL}$ reflects aspects of Hacking’s ‘experimental realism’, which holds that “‘natural kind’ is a name for [that which] helps people do better”.¹⁸ Kinds, according to Hacking and $CToR_{INT-REL}$, are categories whose boundaries are determined by those objective distinctions in reality which are relevant to a community’s interests. Though which particular properties fix membership will vary, in all cases kind membership is determined by agent-independent features of reality. Thus $CToR_{INT-REL}$ fulfils the desideratum of weak realism, (2). All properties ϕ determining kind extension are natural, mind-independent properties. Likewise, if ϕ is a property affecting C ’s ability to make predictions and achieve their aims, then ϕ will be projectable. Since sameness-with-respect-to- ϕ will therefore determine kind membership, (1) will be satisfied.

Essentialism. As was shown in §1, essentialism is derivable from the fact that sameness, as a cross-world relation, determines the extension of a kind in all possible worlds. Since $CToR_{INT-REL}$ retains this cross-world sameness relation, kind terms will continue to rigidly designate. Thus, if, given interests i ,

¹⁸Hacking [8] p. 114.

sameness-with-respect-to- ϕ determines the extension of K , then, if the sample a possesses ϕ , membership of K will necessarily entail possession of ϕ ; ϕ will be an essential property of K . CToR_{INT-REL} differs from standard essentialism in that essential properties are changeable across time and communities. Accordingly, essentialist talk ought to be indexed to temporal place and interests.

Cross-cutting. Under CToR_{INT-REL}, kinds determined by differing interests may overlap, violating (5). However, Khalidi¹⁹ proposes a modification of the principle, on the basis that many plausible kind distinctions, such as ‘acid/base’ and ‘organic/inorganic’ naturally cross-cut. Khalidi claims “their ability to co-exist may be explained by saying that they embody different interests”.²⁰ If we follow Khalidi’s qualification, CToR_{INT-REL} satisfies (5), since no kinds with the same background interests will cross-cut.

Conclusion

CToR_{INT-REL} is simultaneously a semantic and metaphysical account of natural kinds. Semantically, it retains Putnam’s externalism, by providing a means of determining relevant sameness-relations. Metaphysically, CToR_{INT-REL} combines features of both essentialism and constructivism, two positions normally considered oppositional. It adopts weakened versions of both positions, making kind essences conditional upon interests, and holding that, though kind boundaries are relative to agents, they are independent of their mental states. In its solution of the relevant sameness problem from §2, CToR_{INT-REL} treats kinds not as stable, unchanging categories, but as classes determined by natural properties which reflect the interests, aims and practices of a linguistic community. Kinds are formed through the process of agents’ interactions with the world, and represent the way the world imposes itself upon agents’ practices.

References

- [1] Boyd, R. (1991) ‘Realism, Anti-foundationalism and the Enthusiasm for Natural Kinds’, in *Philosophical Studies* 61: 127–48.

¹⁹Khalidi [13].

²⁰Ibid. p. 42.

- [2] Devitt, M. (1981) *Designation*. New York, NY: Columbia University Press.
- [3] Devitt, M. & Sterelny, K (1999) *Language and Reality: Second Edition*. Oxford: Blackwell.
- [4] Dupré, J. (1981) ‘Natural Kinds and Biological Taxa’, in *The Philosophical Review* 90(1): 66–90.
- [5] Ellis, B. (2001) *Scientific Essentialism*. Cambridge: Cambridge University Press.
- [6] Evans, G. (1973) ‘The Causal Theory of Names,’ in *Proceedings of the Aristotelian Society, Supplementary Volumes* 47: 187–208.
- [7] Goodman, N. (1970) ‘Seven Strictures on Sameness’, in *Problems and Projects*. Indianapolis, IN & New York, NY: Bobbs-Merrill Company.
- [8] Hacking, I. (1991) ‘A Tradition of Natural Kinds’, in *Philosophical Studies* 61: 109–126.
- [9] Hacking, I. (1993) ‘Working in a New World: The Taxonomic Solution’, in P. Horwich (ed.), *World Changes: Thomas Kuhn and the Nature of Science*. Cambridge, MA: The MIT Press.
- [10] Hacking, I. (2007) ‘Putnam’s Theory of Natural Kinds and Their Names is not the same as Kripke’s’, in *Principia* 11(1): 1–24.
- [11] Housecroft, C. & Sharpe, A. (2001) *Inorganic Chemistry*. Harlow: Prentice Hall.
- [12] International Astronomical Union, Resolution 5A, 26th General Assembly, (2006). Available at http://www.iau.org/static/resolutions/Resolution_GA26-5-6.pdf
- [13] Khalidi, M.A. (1998) ‘Natural Kinds and Cross-Cutting Categories’, in *The Journal of Philosophy* 95(1): 333–50.
- [14] Kripke, S. (1980) *Naming and Necessity*. Oxford: Blackwell.
- [15] Kuhn, T. (1990) ‘Dubbing and Redubbing: The Vulnerability of Rigid Designation’, in *Minnesota Studies in the Philosophy of Science*. 298–318.
- [16] Kuhn, T. (2000) *The Road since Structure*. Chicago, IL: University of Chicago Press.
- [17] Mellor, D.H. (1977) ‘Natural Kinds’, in *British Journal for the Philosophy of Science* 28: 299–312.

- [18] Mill, J.S. (1868) *A System of Logic*. London: Parker, Son, and Bourn.
- [19] Prohens, J. & Nuez, F. (2008) *Vegetables, Vol. II: Fabaceae, Liliaceae, Solanaceae, and Umbelliferae*. New York, NY: Springer.
- [20] Putnam, H. (1975) 'The Meaning of 'Meaning'', in *Mind, Language and Reality: Vol. II*. Cambridge, Cambridge University Press.
- [21] Quine, W.V.O. (1969) 'Natural Kinds', in *Ontological Relativity and Other Essays*. New York, NY and London: Columbia University Press.
- [22] Salmon, N. (1979) 'How Not to Derive Essentialism from the Theory of Reference', in *Journal of Philosophy* 76: 703–25.
- [23] Sienko, M.J. & Plane, R.A. (1971) *Chemistry*. New York, NY: McGraw-Hill.
- [24] Sterelny, K. (1983) 'Natural Kind Terms', in *Pacific Philosophical Quarterly* 64: 110–25.
- [25] Soames, S. (2007) 'What are Natural Kinds?', in *Philosophical Topics* 35: 329-42.

Verificationism and Relativism: A New Defence of Metaphysics

Atticus Gatsby

At the request of the author, this paper has been published under a pseudonym.

Introduction

In this paper I substantiate a new objection to the verificationist attack on metaphysics. To do this, I: (§1) briefly unpack the attack, showing that it consists of two prongs and outlining the theories which support each prong; (§2) establish that for the first prong to be correct, ‘meaningful’ must be a relational predicate; (§3) establish that if ‘meaningful’ is an indexical predicate, there are multiple readings of the first prong, and all relevant readings significantly weaken the attack; (§4) outline a *prima facie* plausible argument for supposing that ‘meaningful’ is not a relational predicate and that, therefore, the first prong is false; (§5) conclude that either the verificationist objection to metaphysics is weaker than has been supposed, or there is a *prima facie* argument for it being false, and ruminates on some possible objections to my arguments.

1 Unpacking the Verificationist Objection to Metaphysics

The verificationist attack on metaphysics claims that metaphysics attempts the impossible in two different ways:

- (1) Metaphysics attempts to treat as meaningful that which has no meaning,
- (2) Metaphysics attempts to increase our knowledge by *a priori* reasoning.

The first of these claims relies on the verifiability theory of meaning which holds that any meaningful statement is either analytic or is “directly or indirectly verifiable”¹ – where direct or indirect verifiability means something like ‘it is possible to derive an actual or possible observation from the statement’.

¹Ayer [1] p. 17.

The second claim relies on the theory that all *a priori* truths are analytic and therefore cannot increase our knowledge.²

The purpose of this essay is to establish a new argument against (1). In doing so, I will be considering (1) on its own – abstracted from (2). Consequently, my conclusion will not be that verificationism fails; it will be that (1) has little to no force without (2). This conclusion is interesting because it entails that for verificationism to pose a strong problem for metaphysics, it must be (2) that does the work, or (1) must be shown to gain extra support from (2) in a way that has not been previously demonstrated.

2 Meaningfulness and Relativism

A consequence of meaningfulness being conferred by verifiability is that the status of a sentence's meaningfulness is hostage to the novel epistemic situation of whoever is uttering it. This is made clear by asking the question "by whom need an utterance be verifiable?"

Imagine the scenario in which two people (P_1 and P_2) each independently utter the same sentence: 'the Mona Lisa is on display at the Musée du Louvre'. The proposition this sentence expresses is a mundanely verifiable sentence for P_1 who is easily able to go to the Musée du Louvre and see the painting. But P_2 was born blind, deaf and with extensive damage to her nerve endings such that she cannot feel pressure, pain or temperature and is consequently unable to point to any evidence or observation statements which she could use to verify the proposition.³ In this situation, it seems that the verificationist is committed to saying that P_1 and P_2 's separate utterances of exactly the same sentence have different statuses of meaningfulness and, accordingly, that 'meaningful' is a relational predicate.

Some might wish to challenge me on this by pointing to something like the "*division of linguistic labour*".⁴ Such a challenge would claim that when P_2 utters the proposition, it is in some useful sense verifiable by virtue of the fact that P_2 is a member of a linguistic community that contains persons who are able to verify the location of the Mona Lisa.

²Ibid. pp. 21–24 and pp. 103–115.

³Presume that P_2 learnt enough language to understand the utterance before acquiring the maladies that render the proposition unverifiable for her.

⁴Putnam [6] p. 144.

Setting aside questions about whether or not someone in P_2 's unfortunate condition could really be said to be a member of any linguistic community, my thought experiment can easily be amended to meet the spirit of this challenge, with an injection of science fiction. Imagine an alien species on a planet so distant that no member of that species could be called a member of the same linguistic community as any being on this planet. Imagine this alien species is so far away and so technologically impoverished that it is nomologically impossible for any member of that species to make any observation as to the location of the Mona Lisa. Were a member of that species to utter the same sentence (expressing the same proposition) as P_1 and P_2 , it does seem that the verificationist really would be committed to applying different values of meaningfulness to the same proposition. This shows that the verificationist is committed to the view that 'meaningful' is a relational predicate. If 'meaningful' is not a relational predicate, the verificationist cannot make sense of situations like these.

It might be responded that it is *metaphysically* possible (or perhaps 'possible in principle'), for a member of this alien species to make an observation as to the location of the Mona Lisa, and that this is the kind of possibility that the verificationist is talking about. However, this would be a strange reading of verificationism, since it could then be asked: 'what actual or possible observation statement could a member of this alien species point to in order to verify the metaphysical possibility of discerning the location of the Mona Lisa?' In general, I suspect that those who oppose metaphysics are probably not enthused by the prospect of appeals to metaphysical possibility.

Along similar lines, one might object that there is an important counterfactual conditional claim which is true for the alien utterer – and it is because of this that the alien's utterance is meaningful. The counterfactual claim would be something like 'if I were in some situation S, then I would be able to deduce an observation statement O which verifies the sentence 'the Mona Lisa is on display at the Musée du Louvre', but this seems like another strange reading of verificationism because it is not clear what kind of observation could verify this counterfactual claim. As such, I think it more appropriate to take the verificationist as referring to nomological possibility and, as described above, this entails that 'meaningful' is a relational predicate.

3 On the Varieties of Interpretation

If 'meaningful' is an indexical predicate, a degree of ambiguity is introduced into (1), which can now be interpreted in many ways. The way that (1) is inter-

preted partly depends on what meaningfulness is relative to. Some candidates include epistemic states, natural capacities, and Wittgensteinian ‘techniques’. For the purposes of this paper, I will consider epistemic states, but the same arguments can be made *mutatis mutandis* for other interpretations. With this in mind, (1) can now mean the following:

- (1.1) Metaphysics attempts to treat as meaningful that which has no meaning from any epistemic state among those possessed by persons alive today.
- (1.2) Metaphysics attempts to treat as meaningful that which has no meaning from any logically possible epistemic state.

The number of possible interpretations is equal to the number of possible responses to the question “from which epistemic states do these things (which metaphysics attempts to treat as meaningful) have no meaning?” I suspect that this means there are an infinite number of possible interpretations (since there are plausibly an infinite number of possible epistemic states), but only the two interpretations above are relevant to this discussion.⁵

On either of the above interpretations, the verificationist view is weakened. If (1.1), then the verificationist is only saying that metaphysics is meaningless to people alive today, and does not rule out the possibility that metaphysics is meaningful to some person(s) somewhere at some time. This does not render metaphysics possible now, but it is an interesting retreat for the verificationist – particularly interesting if (1.1) is understood in terms of natural capacities, because it then has a distinctly Kantian feel. If (1.2), then the verificationist needs to establish that all metaphysical propositions are such that it is *logically impossible* that any *logically possible* epistemic state would be equipped to derive an observation statement from them. It is not clear how the verificationist could achieve such a gargantuan task. But there are still more problems for the verificationist.

⁵Owing to the fact that variance in meaningfulness is due to variance in epistemic state, any possible interpretation will take the form:

Metaphysics attempts to treat as meaningful that which has no meaning for all persons possessing any epistemic state in the set of epistemic states X.

If X = any set that does not include all currently existing epistemic states, then metaphysics is possible for at least 1 person alive today and the verificationist objection to metaphysics is false.

If X = the set of only all epistemic states possessed by persons alive today, then it is equivalent to (1.1).

If X = the set of all logically possible epistemic states, then it is equivalent to (1.2).

If X = any other set, then the arguments applied to (1.1) or (1.2) will have traction on it.

4 A *Prima Facie* Argument Against the Relativity of Meaningfulness

I now turn to give a *prima facie* plausible argument for thinking that ‘meaningful’ is not an indexical predicate. Considerations of brevity prevent me from attempting a full and proper argument to this effect. The first part of the argument has been appropriated from the context of the debate surrounding token-reflexive theories of A-sentences in the philosophy of time. It is Smith’s argument.⁶

Take a plausibly true sentence like the following:

(A) “It was true that the era devoid of linguistic utterances is present.”

Smith points out that (A) is a “past tensed ascription of a truth value true to the truth vehicle expressed by the clause following the operator ‘it was true that’”. Smith then points out that (A) can only be true if the truth vehicle is not a linguistic utterance, since there cannot possibly be a true utterance of “the era devoid of linguistic utterances is present”. This entails that, if (A) is true (which seems plausible), there must be some truth vehicle which is not a linguistic utterance. “The natural choice of this vehicle would seem to be a proposition”, Smith concludes.

If Smith’s *prima facie* plausible argument is right, then there is some sense in which some things like sentences or propositions exist and have truth values independently of being uttered. Now I wish to employ Smith’s conclusion as part of a new argument:

(P1) Propositions have truth values independent of being uttered (this is Smith’s conclusion).

(P2) Only that which is meaningful can have a truth value.⁷

(C1) Propositions are meaningful independent of being uttered.

(P3) (1) requires the meaningfulness of sentences to depend on the context of their utterance (this is equivalent to ‘meaningful’ being relational).

(C2) (1) is false.

⁶See Smith [7] pp. 72–73.

⁷For brevity, I take this to be uncontroversial.

This is not a watertight or knockdown argument, but it does seem *prima facie* to present a problem for (1). Were I afforded the luxury of a higher word limit, I would expand on this argument considerably.

5 Conclusion and an Objection

In conclusion, I have demonstrated that:

- (a) The first prong of the verificationist attack on metaphysics relies on the verifiability theory of meaning which requires ‘meaningful’ to be an relational predicate.
- (b) If ‘meaningful’ is an indexical predicate, the first prong has many possible readings which either do not render metaphysics *completely* impossible, or render (1) insufficient to cause a problem for metaphysics (but perhaps (2) can be brought in to help).
- (c) There is a *prima facie* plausible argument for thinking that ‘meaningful’ is not an relational predicate and, therefore, that the verifiability theory of meaning is false and, by extension, all interpretations of (1) are false.

I will now briefly respond to a possible objection.

It might be objected that the arguments in §4 beg the question because they employ a metaphysical claim (that there exist strange meaningful entities called sentences or propositions) to defend metaphysics from the charge that precisely those kinds of claims are meaningless.

My response is that the same move is made by the verificationist. If the arguments of §4 really are question-begging, they still ought not to be dismissed because they demonstrate that the verificationist requires the elimination of metaphysics (in this case, the elimination of claims about entities like sentences or propositions) in order to establish the elimination of metaphysics. Hence, to whatever extent the arguments of §4 are flawed, so too is this argument against it. Perhaps this points to a need for some further argument, hitherto not alluded to, in order to settle the debate. Perhaps it points to an impasse.

References

- [1] Ayer, A.J. (1980) *Language, Truth and Logic*. Middlesex: Penguin Books.

- [2] Blanshard, B. (1996) 'In Defense of Metaphysics', in W.E. Kennick & M. Lazerowitz (eds.) *Metaphysics: Readings and Reappraisals*. New Jersey: Prentice-Hall, Inc.
- [3] Church, A. (1949) 'Review of Language, Truth, and Logic', in *Journal of Symbolic Logic* 14: 52–3.
- [4] Hamlyn, D.W. (1984) *Metaphysics*. Cambridge: Cambridge University Press.
- [5] Lazerowitz, M. (1995) *The Structure of Metaphysics*. London: Routledge and Kegan Paul.
- [6] Putnam, H. (1975) 'The Meaning of 'Meaning'', in *Minnesota Studies in the Philosophy of Science* 7: 131–93.
- [7] Smith, Q. (1993) *Language and Time*. Oxford: Oxford University Press.
- [8] Wright, C. (1989) 'The Verification Principle: Another Puncture – Another Patch', in *Mind* 98: 611–22.

A Critique of Jeffrie Murphy on Resentment*

Bjorn Wastvedt

St Olaf College, MN

1 Introduction

Normative talk of resentment falls into two categories: conditional approval and complete disapproval. One of the strongest voices in the former group is Jeffrie Murphy, especially as he presents his arguments in the second chapter of his book *Getting Even: Forgiveness and Its Limits*.¹ Murphy claims to establish the moral legitimacy of the vindictive emotions and of resentment in particular. In what follows, I argue that Murphy's theory of the emotions and his moral commitments do not, in fact, legitimise resentment; he builds upon powerful theoretical considerations only to conclude his argument invalidly in favour of resentment.

To understand the place of resentment in human emotion, I first present an overview of Murphy's views on the roles and nature of reason and emotion in human interaction (§2). I then present Murphy's argument for the legitimacy of resentment (§3). Finally, I object to Murphy's defence, showing the logical invalidity and practical weakness of the main conclusion of his argument (§5). To understand Murphy's position, it is appropriate to start with a central statement from his work.

Vindictive passions can legitimately be attributed to sane and virtuous people. Virtuous people can, I think, sometimes even enjoy without guilt the knowledge that those who have wronged them

*I wish to thank Jeffrie Murphy for graciously offering an extensive and valuable response to the present draft. His suggestions, and Dr Ed Santurri's comments, have greatly advanced my thoughts on resentment and love, as have the helpful comments I have received on the paper from Dr Charles Taliaferro, Patrick J. Behling, Miriam Brown, Elizabeth Lemke, and Maren Beckman. I also gratefully acknowledge my fellow students in Dr Santurri's seminar 'Religion 395: Ethics after the Fall', for which I originally wrote this paper, and my colleagues in the St Olaf chapter of Phi Sigma Tau; in both settings, even my most persistent incoherencies have been met with helpful discussion and advice.

¹Murphy [10] pp. 17–26.

are ‘getting theirs’.²

Below, I will show that the position stated above cannot be defended using Murphy’s arguments.

2 Reason and Emotion: Murphy’s Framework Developed

I first seek to clarify the nature and place of certain emotional reactions in human interaction within Murphy’s psychological framework. Though Murphy does not make a description of such emotional reactions explicit, he does sprinkle relevant comments throughout his chapter on vindictiveness from which we can reconstruct a coherent and representative account of emotional reactions. Murphy’s coauthored *Forgiveness and Mercy* with Jean Hampton also provides helpful bits.³ A formalisation of Murphy’s description of the relevant emotional reactions requires some theoretical structure with which he would agree.

In his chapter of *Getting Even* on vindictiveness, Murphy refers casually to “emotions” and “passions” in the moral realm, beginning with his opening question:

In which category does vindictiveness belong – among those passions that no morally decent person would willingly retain (malice, cruelty, spite, racial hatred, etc.) or among those passions that bring moral credit to the person who possesses them (kindness, generosity, indignation over wrongs done to others, etc.)? Or does it perhaps straddle both of these categories?⁴

Murphy’s question frames much of the following debate; for the moment, I use it only to demonstrate his terminology. We can cast Murphy’s terminology – the many “passions” that he lists here – in a more structured framework by appealing to Peter Strawson, the one philosopher whose positions on emotion Murphy cites in his text.

²Ibid. p. 25; for philosophers supporting Murphy see French [6] and Brudholm [2].

³Murphy & Hampton [9] pp. xi–xii, 1–10, 14–19, 88–110. In general, Murphy’s other works do not clarify his thinking about the role of emotions. The two works of his cited here are by far the best references on the matter.

⁴Murphy [10] p. 18.

Murphy explicitly identifies his moral “emotional responses” or “passions” with Strawson’s “reactive attitudes” in the following passage:

What Peter Strawson calls the “reactive attitude” of resentment, directed toward wrongs and those who do the wrongs, is a paradigm example of such emotional response.⁵

Murphy’s statement here advances the present discussion in a number of ways. First, it allows us to adopt Strawson’s terminology and (later) attendant considerations into Murphy’s theory of emotions. Second, it shows that resentment holds an important place in both philosophers’ theories of the emotions. As resentment is the subject both of Murphy’s argument and of my critique, a theory of the emotions developed in tandem with Strawson will be an appropriate base for my argument. Finally, we can see the centrality of the interpersonal nature of the “reactive attitudes.” Up for consideration are the reactive attitudes, or emotional responses, feelings which occur in the context of community. The notion of community is central as I formulate my positive position in response to Murphy later on in the paper.⁶

Given the considerations above, I adopt Strawson’s term (“reactive attitudes”) to characterise the passions and feelings at issue here.⁷ The reactive attitudes include hatred, compassion, indignation, love, and resentment, to name a few. Though all reactive attitudes are passions, there are certainly passions which are nonreactive. Neither my joy at seeing a blue sky nor my enthusiasm for croquet are reactive attitudes in the relevant sense because neither involves another human party. In my proposal of a replacement to resentment at the end of the paper, this distinction must be kept in mind. Whatever can fill the place of resentment must be a reactive attitude and not merely a non-reactive passion. With these definitions in order, I consider two related characteristics of the reactive attitudes: how one can understand them theoretically and what role they play (along with reason) in human interaction for Murphy.

Murphy’s most direct comment on the characterisation of the reactive passions in general comes at the beginning of his consideration of their morality and rationality. In fact, his argument (for the moral legitimacy of resentment in particular) turns on these two qualities of the reactive passions. His main

⁵Ibid. p. 19.

⁶I believe that it is possible to feel many of the emotions Murphy lists above, even resentment, toward oneself. This case should perhaps be brought up separately, as much of what follows depends on the interpersonal context.

⁷Strawson [13] p. 7.

consideration runs as follows: the reactive passions are either irrational or at least partly rational, and they are either immoral or at least sometimes moral, simply because they are the sort of things that could have these characteristics.⁸ Murphy believes that the reactive attitudes “are at least in part cognitive states, states of belief and not just feeling”⁹ they “involve a component – belief – that is open to rational evaluation”.¹⁰

Given this information, we can again clarify Murphy’s position by identifying him with a school in the philosophy of emotion. I consider Murphy to hold a broadly cognitivist view of emotion; that is, that every reactive attitude (e.g., resentment) entails a believed proposition (perhaps that one has been wronged), though the attitude consists in more than the proposition (in one’s increased irascibility, for example).¹¹ This latter part we might term arational. Murphy’s characterisation of the reactive passions comports rather well with the standard definition of emotional cognitivism.

Having explained Murphy’s general theory of the emotions, I consider Murphy’s account of the role of the reactive attitudes in human interaction. The question at hand is descriptive, not normative. Thus we ask, “What roles do emotion and reason play? Must human action necessarily consist of both? Is it possible to eliminate one or the other?” These considerations play an essential role in my argument that follows.

Before proceeding to Murphy’s views on the matter, I first wish to show that our intuitions about the roles of reason and emotion in human behaviour can be confused and contradictory. Often, our reason appeals to our morality to criticise the presence of the reactive attitudes, especially in their extremes; yet these same attitudes seem to ground human relationships. We denounce the holding of grudges, for example, but praise proper indignation (whatever that may be). Aristotle advocated a mean between excessive irascibility and “inirascibility,” while the cognitivist Stoics thought that all emotion was irrational (and therefore inadvisable). Finally, as mentioned above, the propositional content of a reactive attitude (a belief) is rationally analysable, but the side-effects it entails are not as directly approachable. To investigate the reactive attitudes, we must keep both aspects of the reaction in mind. The side-effects,

⁸Murphy also gives empirical support for these claims about the reactive passions in terms of rationality and morality in the same chapter.

⁹Murphy & Hampton [9] p. 5.

¹⁰Murphy [10] p. 22.

¹¹De Souza [4]; for a classic defence of emotional cognitivism, see Solomon [12] pp. 251–81.

or the arational component, make it what it is, *viz.* a reactive attitude; but we will analyse its rationality and morality – at least in an internal or principled sense – in virtue of its propositional content.¹²

We see below that Strawson argues for the brute necessity of reactive attitudes in human interaction, though the reactive attitudes themselves may not be rational. If Murphy wants to follow Strawson on this point, he does not succeed, for the characterisations that Murphy makes of the role of the reactive attitudes actually point away from the necessity of the reactive attitudes. Murphy does not address the issue directly; however, several of his passages are especially pertinent, especially a sentence from *Getting Even*:

Intellectually believing something and actually feeling it in your guts emotionally are, of course, two quite distinct things – although we like to deceive ourselves into confusing them.¹³

These words seem to establish a dichotomy – or possibly a spectrum – in the existence of the reactive attitudes and emotion. On the one hand, reactive attitudes can be entirely absent when one is just “intellectually believing something.” On the other hand, reactive attitudes can accompany such belief (no doubt by means of the attitudes’ propositional components). Murphy goes on to draw conclusions about what the possibilities of the presence or absence of the reactive attitudes means about the agent’s conviction in his or her belief.¹⁴

Certainly, Murphy is not so naïve as to think that pure intellectual belief actually happens very often, especially in the moment when one feels wronged by another party. However, the very fact that he makes a distinction between cases of unemotional commitment and cases of emotional commitment to moral propositions shows that he does not think that the reactive attitudes are necessarily a part of certain beliefs, and that it is likely not just in some sort of limit-case scenario that the two are separated. This conclusion plays an important role in what he has to argue: not only must he show that resentment is a morally justifiable reaction among other possible emotional reactions to a situation, but he must show that it is not better to aim (even futilely) to

¹²Of course, there are other ways to consider a reactive attitude in a moral light, depending on the moral framework one wishes to use. Murphy’s moral considerations will become clear presently.

¹³Murphy [10].

¹⁴I realise that the claim in this paragraph is somewhat controversial, especially given Strawson and Butler’s opposite views on the matter; however, I do not see any way in which Murphy’s text could be postulating the necessary coexistence of reason and emotion.

eliminate the reactive attitudes as much as possible (as the Stoics would have us do), for under his view, the latter option is indeed at least a coherent possibility. Having understood Murphy's thoughts on the status of the reactive attitudes in general, I turn to Murphy's argument in support of resentment.

3 Murphy's Argument for Resentment

Murphy begins his argument in the second chapter of *Getting Even* by considering vindictiveness. He defines vindictive emotions as "harsh negative passions – anger, resentment, even hatred – often felt by victims toward those who have wronged them".¹⁵ Thus Murphy would agree that vindictive attitudes are reactive attitudes, and his theory of emotion described above applies here. Vindictiveness can, as Murphy states, mean any number of things, from anger to resentment; here I consider just his argument for the particular attitude of resentment.

Murphy begins by considering the possibility of the rationality and morality of resentment as mentioned above in his description of the reactive attitudes. He points out that if resentment was either irrational or immoral, one could not defend it in any circumstances. Few, it seems, would disagree with such a claim. To rule out the possibility of the irrationality or the immorality of resentment, he must show that resentment is rational and moral. Murphy's argument for the rationality and morality of resentment proceeds in two parts: First, he attempts to show that resentment is necessary to fulfil certain fundamental values we would not be willing to sacrifice. Second, he advances several miscellaneous arguments designed to show the reasonableness of resentment.

In the first part of his argument (which is also where the argument's main thrust lies), Murphy begins by stating that reactive attitudes in general are necessary components of strong moral commitment. If one is strongly committed to a moral principle, he holds, one should feel not only intellectual allegiance to it but one should experience reactive attitudes in relation to it:

The virtuous person will not simply say and believe in a purely intellectual way that he respects himself as a free and equal moral being with basic rights; he will also react emotionally if he is not treated as such a being.¹⁶

¹⁵Ibid. p. 17.

¹⁶Ibid. p. 19.

Virtue, at least virtue in a sinful, fallen world, entails some sort of moral emotion.¹⁷

Murphy next identifies two non-negotiable, positive values: self-respect and respect for the moral order.¹⁸ As Murphy does, we shall consider these two values simultaneously. Working from the premise in his description of the reactive attitudes above that strong moral commitment actually involves reactive attitudes, Murphy advances his main argument in support of resentment in one key passage:

Resentment stand[s] as emotional testimony that we care about ourselves and our rights. This passion . . . also stands as testimony to our allegiance to the moral order itself. . . . If we do not show some resentment to those who, in victimising us, flout those understandings, then we run the risk – in Aurel Kolnai’s words – of being “complicitous in evil”.¹⁹

Murphy makes three claims in this excerpt, each in support of the morality of resentment.

First, he claims that (1) resentment evinces strong moral commitment to our own rights. He then asserts that (2) resentment also evidences strong moral commitment to our moral principles. His third claim is the most important. (3) The lack of resentment, he claims, risks being “complicitous in evil.” One can read this remark in at least three ways. Complicity here implies assent to an evil act, or at the very least a lack of dissent to it. So first, perhaps Murphy actually means that those who lack resentment risk actually cooperating in the evil done.²⁰ More conservatively, Murphy may mean that those who lack resentment show an indifference to the evil at hand. I only attribute to Murphy an even weaker claim: that those who lack resentment show themselves to be like Murphy’s purely intellectual moralist, the man who has no emotional commitment (and thus for Murphy, a good deal less commitment generally) to the moral principle at hand. Murphy considers these three points, taken together, to be a strong argument for the moral legitimacy of resentment.²¹

¹⁷In this model, what would the prelapsarian state of emotion be?

¹⁸Ibid. p. 19.

¹⁹Ibid. p. 20.

²⁰As we see in the penultimate section, this notion is not so far-fetched, though I do not believe that Murphy would assent to it.

²¹For another defence of resentment, see Thomas [14] pp. 211ff.

Notably, he is not alone in his conclusion: Aurel Kolnai, whom he quotes above, gives a staunch defence of resentment:

The retributive attitude as such is nothing but the correct primary response to immoral conduct as such, and above all to ‘wronging,’ i.e., immoral conduct specifically hurtful or offensive to the person ‘concerned,’ the primary claimant to retribution.²²

In the second part of his argument, Murphy offers other ideas in support of resentment. He claims self-defence as a third non-negotiable value of the type mentioned above. Murphy’s argument from self-defence is, as he allows, an “instrumental point”: he claims that by sacrificing resentment, we open ourselves to prolonged attack.²³ Murphy also wants to claim here the efficacy of resentment institutionalised in the judicial and prison system, for without the negative reactive passions (he implies), those measures would not be in place. He also hails Bishop Joseph Butler for providing perhaps the first modern defence of resentment, that as a God-given, natural emotional response, resentment must serve some useful end.²⁴ A similar defence could be made, he says, without theology by appealing to evolutionary biology. He also argues with Bishop Butler that our negative reactive attitudes (like resentment) help us to recognise right and wrong.²⁵ Needless to say, numerous arguments of a similar kind are advanced in contemporary literature on vindictiveness.²⁶ We must differentiate these points of secondary importance from Murphy’s main argument.

Having marshalled all he can in support of resentment, Murphy concludes his defence of the attitude with a challenge to his opponents, which he implies must be answered in order to defeat his position.

It may be reasonable to impose – for a change – the burden of proof on those who would challenge beliefs in [resentment’s] legitimacy – impose on them the burden of actually *arguing* for

²²Kolnai [8] p. 93.

²³Murphy [10] pp. 19–20; see also Brudholm [2] p. 150 and Améry [1]. I take Jean Améry’s general position (that his resentment is a historical-cultural duty to the post-Holocaust generations) to be a form of the same argument, but on a much more serious scale.

²⁴Murphy [10] p. 18; Butler [3] pp. 90–102. Butler specifies several useful ends for resentment; among them several that Murphy picks up.

²⁵Murphy [10] p. 18; Butler [3] pp. 90–102.

²⁶For example, French [6] and Brudholm [2].

the irrationality or immorality of these emotions, and not merely trotting out sentimental clichés.²⁷

Murphy considers several objections which address this challenge: Nietzsche on the irrationality of resentment, Karen Horney on consequentialist worries about the immorality of resentment, and Jean Hampton on the ineffectiveness of resentment.²⁸ Murphy competently responds to each objection, and for the sake of argument, I make the not altogether implausible assumption that he could likely respond in a similar fashion to other attempts to answer the challenge. Murphy seems to believe so, too, as he ends his account of vindictiveness on quite a positive note:

I think it is justified to conclude – at least provisionally – that vindictive passions can legitimately be attributed to sane and virtuous people. Virtuous people can, I think, sometimes even enjoy without guilt the knowledge that those who have wronged them are “getting theirs”.²⁹

By his conclusion, Murphy claims that in certain circumstances the presence of the vindictive attitudes contributes to the subject’s virtue. Resentment can be the right thing to feel.

4 A Critique of Murphy on Resentment

My critique assumes Murphy’s theory of the emotions and of resentment and addresses Murphy on his main point, leaving the second part of his argument alone. Though the ideas brought forth in that latter portion of his argument do carry weight in his defence of resentment, it is fair to say at this point that without the strength of his main argument, Murphy would not be comfortable in drawing his conclusion.

This description of my critique leads to another important characteristic of what I have to say. As discussed above, Murphy sees as central to the issue the rational and moral status of resentment: he thinks that either it is the case that resentment is sometimes rational and sometimes not immoral or it is the case that resentment is in principle irrational or immoral (or both). Hence he

²⁷Murphy [10] p. 22.

²⁸Ibid. pp. 22–25.

²⁹Ibid. p. 25.

raises the challenge quoted above, namely that to argue against his position, one must argue for either the irrationality or the immorality of the vindictive passions in principle. This predicament is a dilemma to be sure: one horn involves agreeing with Murphy on the contingent morality of resentment, and the other horn involves showing that resentment is immoral in principle. I wish to accept Murphy's description of resentment and his characterisation of it as at least partly rational, and I do not wish to argue that resentment is immoral. In other words, I aim to go between the two unacceptable alternatives that Murphy opposes. I show that, in Murphy's succinct statement of his argument quoted above, Murphy can draw certain conclusions about the legitimacy of the emotions, but he cannot extend those valid conclusions to resentment with a respectable force of argument.

Murphy's theory of the reactive attitudes enables us to expose his challenge as unfair and his conclusion as not well-supported (though, for the moment, not necessarily false). It allows us to see a way out of Murphy's challenge as stated above: we can respond that resentment is partly rational and may possibly be moral in some circumstances, but we will add that Murphy's commitment to the appropriateness of resentment when wronged overstates his case. All his theory of the reactive passions can show is that some emotional response must be present in moral commitment.

Murphy's crucial mistake lies in the three conclusions he draws in the first part of his argument for resentment. These are that resentment evidences strong moral commitment to our own rights and to our moral principles (from 1 and 2), and that those who lack resentment show the weakness of their moral commitments to be that of Murphy's purely intellectual moralist (from my interpretation of 3). In the analysis of reactive attitudes above, I agreed with Murphy that reactive attitudes are necessary components of a strong moral commitment. Thus (1) and (2) follow: resentment, as a reactive attitude, can validate the strength of our moral commitment. However, conclusion (3) does not follow. The jump in reasoning from concluding that resentment, as a reactive attitude, evinces strong moral commitment to concluding that a lack of resentment betrays a weak moral commitment is the jump from a proposition to its inverse, certainly not always a valid inference. In the argument at hand, Murphy offers no support for conclusion (3) beyond his support for conclusions (1) and (2).³⁰ He offers no reason, in other words, why resentment is

³⁰What I have been referring to as the second half of Murphy's argument for resentment, that is, his secondary and more numerous considerations, also do not seem to support his conclusion in (3). It would be rewarding to investigate possible arguments for identifying resentment as the

the only reactive attitude which can validate our strong moral commitment to principles violated by the thief, the murderer, or the torturer.³¹

Without conclusion (3), Murphy's main argument falls apart. Conclusions (1) and (2) do not establish the necessity or the morality of resentment; they merely agree with Murphy's descriptive account of resentment. The proposition that Murphy needs to defend is not just that reactive attitudes are important to hold if one has a strong moral commitment, but that the specific reactive attitude of resentment is the right one to hold at least in some circumstances. This crucial proposition is Murphy's conclusion, (3); unfortunately, he seems to infer it from conclusions (1) and (2). He cannot do this without very good reasons to consider why resentment should fill the place in our passion for our moral convictions.

The most that Murphy can conclude is that resentment is "a testimony to our allegiance to the moral order" and "a testimony of self-respect".³² To assert (3) confidently, Murphy would need the additional assumption that resentment is the only appropriate reactive attitude to satisfy the needs of (1) and (2). In other words, he would need to show that other reactive attitudes could not work in the place of resentment.

5 Conclusion

Contemporary debate on the issue of resentment does not deal in absolutes. Above, I have offered counter-arguments to Murphy's main argument in favour of resentment. As the conversation on reactive emotion and resentment continues, the considerations offered above expose the invalidity of an influential argument in favour of resentment. The burden of proof in the resentment debate now lies more with those who would assert the moral legitimacy of resentment.³³

sole possible fulfilment of the criteria behind (1) and (2).

³¹See also Murphy [11] and Dostoevsky [5]. Dostoevsky gives a paradigmatic example of a case which might seem to prompt justifiable resentment. See, cited, Murphy's discussion of forgiveness with respect to Dostoevsky's "torturer-general".

³²Murphy [10] p. 20.

References

- [1] Améry, J. (1998) *At the Mind's Limits: Contemplations by a Survivor On Auschwitz and Its Realities*. Bloomington, IN: Indiana University Press.
- [2] Brudholm, T. (2009) *Resentment's Virtue: Jean Amery and the Refusal to Forgive*. Philadelphia, PA: Temple University Press.
- [3] Butler, J. (2006) *The Works of Bishop Butler*, D. E. White (ed.). Rochester, NY: University of Rochester Press.
- [4] de Sousa, R. (2010) 'Emotion', in E.N. Zalta (ed.) *Stanford Encyclopedia of Philosophy* [Online] Available at <http://plato.stanford.edu/entries/emotion/>
- [5] Dostoevsky, F. (1993) *The Grand Inquisitor: with Related Chapters from The Brothers Karamazov*, C. B. Guignon (ed.). Indianapolis, IN: Hackett Publishing Company.
- [6] French, P.A. (2001) *The Virtues of Vengeance*. Lawrence, KS: University Press Of Kansas.
- [7] King, M.L. & Jr. (1992) *I Have a Dream: Writings and Speeches That Changed the World*, J.M. Washington (ed.). San Francisco, CA: Harper-One.
- [8] Kolnai, A. (1973–74) 'Forgiveness', in *Proceedings of the Aristotelian Society*, New Series, 74: 91–106.
- [9] Murphy, J.G. & Hampton, J. (1988) *Forgiveness and Mercy*. Cambridge: Cambridge University Press.

³³I have recently communicated with Jeffrie Murphy about the arguments in this paper. Murphy suggests that resentment can actually evince respect for the other party.

Resenting injuries is a way of showing respect for the wrongdoer as an autonomous rational responsible agent – and not someone who is, for example, not responsible because mentally ill. So in a sense it might be that resenting a person . . . is to pay that person a kind of compliment. (personal communication, 6 April 2012)

One idea to replace resentment is to use the sort of active love advocated by Zossima in *The Brothers Karamazov*. In that case, Murphy's suggestion can be applied inversely as a critique as well. One could worry that reactive love could be taken as resentful after all; certainly the power of an unlooked-for, obstinate, or self-righteous action is not to be underestimated.

- [10] Murphy, J.G. (2003) *Getting Even: Forgiveness and Its Limits*. New York, NY: Oxford University Press.
- [11] Murphy, J.G. (2009) 'The Case of Dostoevsky's General: Some Ruminations on Forgiving the Unforgivable', in *The Monist* 92(4): 556–82.
- [12] Solomon, R.C. (1980) 'Emotions and Choice', in A.O. Rorty (ed.) *Explaining Emotions*. Berkeley, CA: University of California Press.
- [13] Strawson, P.F. (2008) *Freedom and Resentment and Other Essays*. New York, NY: Routledge.
- [14] Thomas, L. (2003) 'Forgiving the Unforgivable?', in E. Garrard & G. Scarre (eds.) *Moral Philosophy and the Holocaust*. Aldershot: Ashgate Publishing Ltd.

Tackling the Repugnant Conclusion*

Helen Weir

University of Glasgow

1 Introduction

When we are considering how our actions affect future generations one of the most important questions we must ask ourselves is: Ideally, how many people should there be? Derek Parfit has identified that if we try and answer this question using some intuitive steps we are led to an extremely counter-intuitive conclusion; that a very high population with a very low quality of life could be the optimum population. It seems that we have no choice but to accept this conclusion or reject the influential moral views that lead us to it. Within this article I will put forward my own novel solution to the Repugnant Conclusion which I have created by adapting the Critical-Level Principle to create a Neutral-Range. I will demonstrate that, by appealing to a Neutral-Range, we can reduce the counter-intuitive consequences of the Critical-Level Principle. The Neutral-Range is, therefore, a better solution to the Repugnant Conclusion.

2 The Repugnant Conclusion

The most effective example Parfit employs to demonstrate how our ethical intuitions lead us to the Repugnant Conclusion is the Mere Addition Paradox. Parfit suggests that we imagine that there exists a World A, a world that contains a relatively large population in which everyone has a high quality of life. World A then has a group of new people added to it. It becomes World A+. Due to a naturally occurring inequality – rather than an act of injustice – this new group has slightly lower average welfare than the original group in World A. For example, imagine that the second group of people suffer from a mild genetic disability, such as slight near sightedness. The addition of the A+ group will cause the average welfare to drop slightly, but will have no detrimental

*Delivered at the BUPA Autumn Conference 2013 on 23–24 November at the University of Manchester.

effect on the first group (see Fig. 1). Parfit then asks us to consider whether

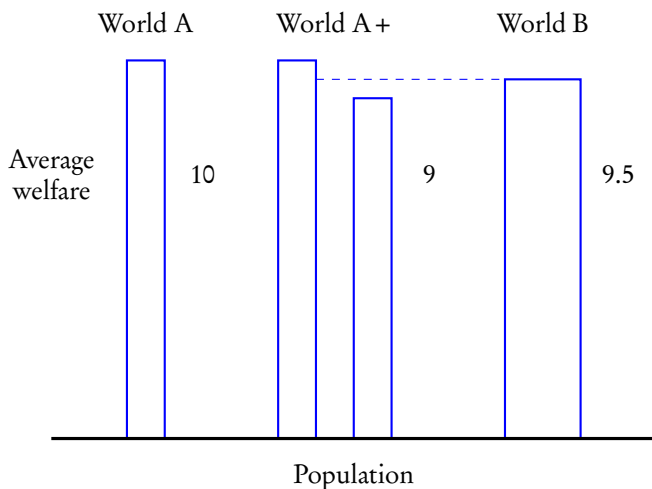


Figure 1: Worlds, A, A+ and B.

World A+ is better than World A. Despite this drop in welfare, I think many of us would not want to claim that A+ is worse than A.¹ This is because, in saying A is better than A+, we are implying that it would have been better if this second group of people had never existed. As their existence has no negative impact on the first A group, and their lives are well worth living, this claim seems counter-intuitive. Moreover, it appears that, despite the overall drop in average welfare, we should claim that World A+ is better than A. World A+ contains double the amount of people living at quite a high level of welfare. Even if we believe that we never have a moral obligation to create more people, I would suggest that almost all of us would agree that World A+ is better than World A.

Parfit then asks us to consider a situation where A+ becomes B.² World B is a population in which the welfare levels of the two groups have been made equal. This move involves a minor gain for the second group at the expense of a little loss from the first group. Furthermore, in B a few more people are added to the population than in A+ and, therefore, the total well-being is higher than either

¹Parfit [4] p. 432.

²Ibid. p. 425.

A or A+. Imagine that in population B an operation to cure the disability of those in the second group has been found. This operation has been paid for by the better off group and the birth-rate has increased. It is hard to resist the claim that B is better than A+. Firstly the total welfare of the population is higher than in A and A+. In addition, if we have any egalitarian principles this move will appeal to them. In accordance to the Rawlsian Maximin, the worse-off group is bettered at the expense of a minor loss from the best-off group. Unless we hold elitist views, and believe that the welfare of the best-off group is more important, we must conclude that World B is better than World A+.³

Therefore we come to the conclusion that A+ is better than A, and B is better than A+. By the law of transitivity B must be better than A. This does not seem to be a too controversial conclusion. In B there is much higher total welfare, and only a little less average welfare, than in A. But, if B is better than A, then presumably if we repeated these steps C would be better than B. Suppose another group with slightly lower welfare is added on to B, creating B+. The welfare of these two groups is then made equal, B+ becomes C. C will be better than B; B is better than A. Therefore, C must be better than A. If we repeat these steps indefinitely eventually we will reach a population that has much more total welfare than A, but has a much lower average welfare. Through a series of inoffensive manoeuvres we have once again reached the conclusion that Z is better than A. It seems that the only way we can avoid this conclusion is by either insisting that A+ is worse than A, that it would have been better if people whose lives are worth living had never existed, or that A+ is better than B, and commit to an anti-egalitarian viewpoint.⁴

3 Conclusion with Even More Repugnance

Within this section I will focus on one particular solution put forward to solve the Repugnant Conclusion, the Critical-Level Principle. I have chosen to focus on the Critical-Level Principle because it is a solution that succeeds in avoiding the Repugnant Conclusion while still allowing us to accept the intuitions which lead us into the Mere Addition Paradox. Many solutions, such as an appeal to Perfectionism⁵ which lead to the anti-egalitarian Elitist Conclusion,

³Ibid. p. 426.

⁴Parfit [4] p. 428.

⁵Parfit [5] p. 18.

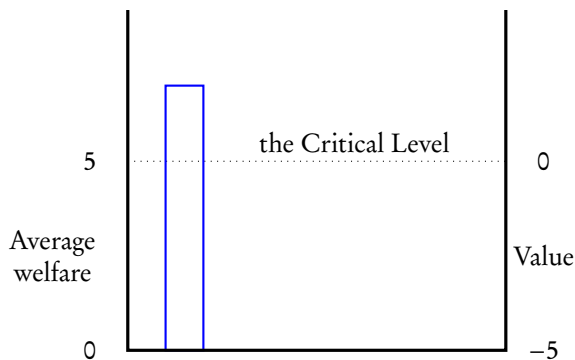


Figure 2: The Critical Level

and an appeal to the Average Principle, which ignores our intuitions concerning the moral impact of numbers, would block the Mere Addition Paradox before it reached Population B. I will not discuss these arguments in details as I do not intend to irredeemably refute them. The purpose of this article is to propose my own solution to the Repugnant Conclusion, not to render all other solutions completely untenable. Before I go on to adapt the Critical-Level approach I will demonstrate why this solution in its current form fails to solve the paradox as it imposes further counter-intuitive consequences.

This version of the Critical-Level Principle is put forward by Blackorby, Bossert and Donaldson.⁶ The Critical-Level is a point on the scale of individual welfare. If a person's welfare level lies below this point their addition to a population will add a negative value (see Fig. 2). However, it should be highlighted that we would not be justified in disregarding a persons' life that falls below the Critical-Level in other moral decisions. If this life was free from suffering, it would still be a life worth living and would have a personal value to the person who lives it. To attain the objective value of a person's life we would need to establish their total welfare minus the Critical-Level.⁷ For example, if a population has a Critical-Level of 5, and a person living within that population has a welfare level of 4, their value will be -1 ($4 - 5 = -1$). The overall value of a population would be calculated by summing up the differences between the Critical-Level and the welfare of all the individuals within the population. A

⁶Blackorby *et al.* [2] pp. 49–50.

⁷Ibid. p. 50.

life that is lived under the Critical-Level will always add a negative value to a population. Therefore, assuming the Critical-Level exists at a level higher than the welfare of the people in population Z, the Repugnant Conclusion will be avoided. Moreover, assuming that the Critical-Level is not placed too high on the scale of welfare the moves from World A to World B in the Mere Addition Paradox can be condoned.

However, one problem with deciding where to put the Critical-Level is that if the Critical-Level starts on the scale of positive welfare, at a point above where a person's average welfare entails they live a life barely worth living, we will be led to the Sadistic Conclusion. The Sadistic Conclusion arises when adding a few people who are suffering to a population leads to a better outcome than adding happy people.⁸ The Critical-Level approach implies the Sadistic Conclusion because we take lives that exist under the Critical-Level to impact negatively to an outcome, even if these lives have a positive average welfare. For example, if you have a Critical-Level of 5, adding 100 individuals whose welfare is at level 4 will reduce the existing 'quality' of the distribution by $100 \times -1 = -100$. If we add 10 individuals who have a negative welfare of -1 , then the 'quality' of the distribution is only reduced by $10 \times -6 = -60$. The second distribution, which contains an additional number of individuals who have lives not worth living, appears to be better than the first, which contains an additional number of individuals whose lives have a positive, if low, value. Therefore, when we are using the Critical-Level Principle, sometimes adding a small number of individuals who are suffering to a situation would be better than adding numerous amounts of people who have lives worth living, though exist at welfare levels under the Critical-Level. Furthermore, this could imply that for a World full of tormented people, World $-Z$, there is a world full of people with lives barely worth living that is worse than $-Z$, provided that this World Z contained enough people.⁹ If we are to accept the Critical-Level Principle in its current form we must reject our intuitions regarding the creation of people who are suffering.

4 The Neutral Range

I will now put forward an adaptation of the Critical-Level Principle that avoids the Sadistic Conclusion. As I have already discussed, Critical-Level Principles

⁸Arrhenius [1] p. 253.

⁹Blackorby *et al.* [2] p. 52.

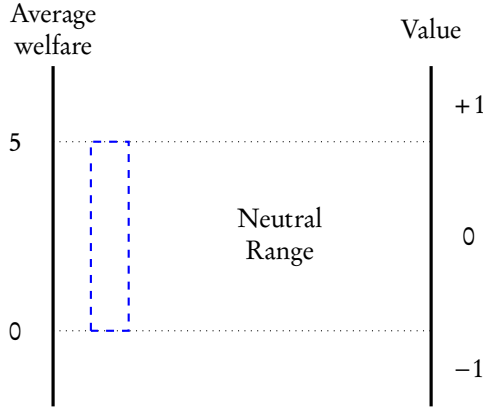


Figure 3: The Neutral Range

which count anyone who has a welfare level which falls under the Critical-Level as negative will either lead to the Repugnant or Sadistic Conclusion. What conclusion arises depends on where the Critical-Level is set. After my objections to both versions of the Critical-Level Principle, it may appear that we should now abandon the Critical Level Principle altogether. However, if we amend the Critical-Level approach very slightly we can avoid almost all of its counter-intuitive import. Instead of claiming that all lives that fall below the Critical Level are negative, we could, instead, appeal to a neutral range. In his article ‘Why Should We Value Population’, Broome suggests that additional lives should add no value to a population, unless they contain extremely high, or extremely low, levels of welfare.¹⁰ Broome does not suggest the Neutral-Range to solve the Repugnant Conclusion but to clarify the moral implications of population growth. He suggests the Neutral-Range would correspond to our intuitions that a life only counts positively to a situation when a person has a very high well-being and negatively to a situation when they have very low well-being. I will take this view of additional lives and apply it in order to solve the Repugnant Conclusion. I will suggest that a neutral range should be an area that exists between a welfare level of 0 (at the point where a life is not worth ending) and wherever we choose to place the Critical-Level (see Fig. 3). If a person has a welfare level which is neither above this range, nor below this range, their life will have an ethically neutral value. In other words, adding

¹⁰Broome [3].

this person to a population will not affect that population's value. However, if a person's welfare is at a level that falls below welfare level 0, or above the Critical-Level, their life would add a negative, or positive, value to a population. Appealing to a Neutral Range, instead of a Neutral Critical-Level, means we will no longer be faced with examples where one hundred people who are suffering have the same value as a thousand people who have lives worth living but whose well-being is just under the Critical-Level. As those who have a welfare level under 0 will have a negative value, the Neutral-Range lets us attribute the value to suffering that our intuitions suggest it should have.

To sum up, as well as recognising the negative value of suffering, the Neutral Range manages to block the Sadistic Conclusion and the Repugnant Conclusion. The Repugnant Conclusion is blocked if the Neutral Range is indexed before a life is barely worth living. This is because, once people reach the Neutral-Range, their addition will add no value to the population. Moreover, the Sadistic Conclusion is avoided completely. It is avoided because any additional people who have a negative level of welfare will always add a negative value to a population. In comparison, any additional people who have a positive welfare level will have a neutral, or positive, impact on a population, depending on whether or not their well-being is above the Critical-level. In other words, if the Neutral Range starts at welfare level 5, it will never be better to add two people of welfare level -1 than any amount that exist within the Neutral Range ($10 \times 0 = 0$ and $2 \times -1 = -2$). The neutral range has managed to block the Repugnant Conclusion and the Sadistic Conclusion without incurring any of the counter-intuitive consequences we have come across so far.

This is not to say that by employing the Neutral Range I have avoided any counter-intuitive consequences. Broome himself rejects it because he suggests a Neutral-Range causes neutrality to be "greedy".¹¹ He claims that adding people who exist within a Neutral Range to a population can cancel out other fluctuations in welfare, fluctuations which we would usually feel compelled to attach moral value to. The Neutral-Range, in the essence of the word, is not neutral.

To demonstrate this objection, Broome asks us to imagine Population A. In this population there exist four people. A woman, let's call her Simone, is now added to this Population A, and it becomes Population B.¹² Simone has a wel-

¹¹Broome [3] p. 407.

¹²Ibid. p. 408.

fare that exists within the Neutral-Range. Therefore, the addition of Simone adds no value to Population B. Broome concludes, that if we are appealing to the Neutral Range, Population A is no better, or worse, than Population B.

Broome then asks us to compare Population B with Population C.¹³ In Population C there are two differences. Simone's welfare is still at a level within the Neutral-Range, but she has been added at a significantly increased welfare level. However, one of the original four people, we will call him John Paul, has suffered a slight decline in welfare. He is now at a welfare level just above the Neutral-Range, when before he existed well above it. For example, in population C John has become slightly hard of hearing and Simone has finally found the right prescription for her poor eyesight. Broome asks us to imagine that the gain in Simone's welfare is enough to outweigh John's loss. He suggests that this would make Population C better than Population B.¹⁴ If Population A is no worse than Population B, Population C is better than Population B, so Population C should be better than Population A.

However, when if we compare Population A with Population C in detail we find that there is little evidence to suggest that the latter is any better than the former.¹⁵ The moves that take us from population A to population C involve, first, the addition of one neutral person, and, second, a decrease in the welfare of another. This is because the addition of a person who exists within the Neutral-Range adds no value to a population. Despite her increase in welfare, within both scenarios Simone's welfare is not above this range. In going from A to C, we have had one bad thing, and one neutral thing happen. A neutral action does not cancel out a negative action. Therefore, it seems that we should evaluate C as worse than A. However, we have already established that population C is better than population B so, by the law of transitivity, this cannot be correct. Broome concludes that neutrality has been greedy, the addition of a neutrally valued person has somehow eradicated the loss of another.¹⁶

However, Broome's claim that the Neutral-Range fails to be neutral can be disputed if we disagree with his original assumption, that population C can be considered no worse than population B. If we decide the Population C

¹³Ibid. p. 408.

¹⁴Ibid. p. 408.

¹⁵Ibid. p. 409.

¹⁶Ibid. p. 409.

has to be worse than Population B, then the rise in Simone's welfare cannot outweigh the loss to John. It is, therefore, not possible for Broome to claim that the addition of a neutral person has cancelled out the decrease in John's welfare during the moves from Population A to Population C. Population C has to be worse than Population B because, by accepting the Neutral-Range, we have already committed to the fact that any level of welfare within the neutral range will add no value to the outcome. Simone's welfare may have increased during the moves from Population B to Population C, even to the extent that we would usually consider it to outweigh John's loss, but Simone's welfare is still within the neutral-range.¹⁷ The increase in Simone's welfare may be of personal value to her, but by subscribing to the Neutral-Range we have already decided it adds no value to an outcome. Simone's welfare level is still objectively valueless. The only change in value we need to take account in the move from Population B to Population C is the decrease in John's welfare. Therefore, under the Neutral-Range, Population C is worse than Population B. Broome's claim that the Neutral-Range is not neutral is ungrounded.

However, I will acknowledge that many of us will find it troubling that, although there are varying levels of well-being within the neutral range, we are choosing to judge all of these levels as equally valueless. If Simone's welfare is at a level of 3 in Population B, and then at welfare level of 5 in Population C, it appears we should take the increase of her welfare into account when comparing the two. We intuitively feel any change in person's welfare, independent of what range that fluctuation falls within, should contribute to the value of a situation. Even if this change of value should only be interpreted as being of personal value, if the only difference between Population B and Population C was that one person had a personal gain, it seems that Population C should be the best population. At this point in my argument we do not have any reason for suspending our appreciation of welfare within the Neutral-Range. This opens my argument up to objections that the application of the Neutral-Range is an arbitrary manoeuvre that just happens to avoid the Repugnant Conclusion.

¹⁷If Simone's welfare had increased to the point she was no longer within the Neutral-Range her welfare would now contribute positively to the population. However that does not mean from the moves from A to C we have an additional neutral person who cancels out a harm done to someone else. Rather, we have originally performed an addition that was neutral but, due to an increase in welfare, this addition now adds a positive value to the situation.

5 Why the Neutral-Range?

In this section I will attempt to refute the accusation that my argument is entirely *ad hoc* by providing a possible justification for the Neutral-Range. I will show that, if we subscribe to a pluralist view of what is valuable, we could be able to explain why we pay attention to welfare when people exist above or below the Neutral-Range, but ignore it when the Neutral-Range is appealed to.

I believe that we could defend the Neutral-Range against the objection that it is arbitrary if we can appeal to pluralism. If we index the Neutral Range using something we value other than welfare – something that we also believe makes a life worth living – this might offer an explanation to why the levels of welfare a life might have within the Neutral-Range are varying but we attribute each life within it a neutral value. Although the lives within the Neutral-Range may be attributed different levels of welfare, we can measure the value of these lives by appealing to something else. We attribute these lives a neutral value because they do not contain something that is also necessary for a worthwhile life.

What we index the Neutral-Range with will depend on what we value in our lives, aside from welfare. Concepts like Autonomy and Success are things that philosophers have often claimed we value intrinsically; we value these things apart from the impact they have on our well-being. A philosopher could claim that a person exists within the Neutral-Range when they no longer have access to these concepts. For example, when a person is autonomous we could count their welfare as attributing positively to a moral outcome. When a person's life is no longer autonomous it will count as contributing nothing to a moral outcome. Finally, a person's life will count negatively when it is both non-autonomous and contains suffering.

However, an argument which specifies concepts we should value outside welfare is bound to be controversial. To defend a pluralist view of the Neutral-Range as a solution to the Repugnant Conclusion, we would not only need to decide on a suitable concept we value unanimously, but we would need to ensure that the presence of this concept diminishes in correlation with welfare.

For now, I will conclude that I have not found a convincing explanation for the Neutral-Range within this article. I will speculatively suggest that an explanation for the Neutral-Range will involve an appeal to pluralism. Identifying what this pluralism will involve is a task I will leave open to other philosophers.

In my final comment I will point out that, even if we are unable to find a jus-

tification for ignoring the different levels of welfare within the Neutral-Range, this does not leave the Neutral-Range at any disadvantage when compared to the Critical-Level Principle. When we apply the Critical-Level Principle we also ignore the value that we would usually intuitively attribute to welfare within a certain range. In fact, we actually distort the value of those who have lives worth living but exist under the Critical-Level. We apply a negative value to their positive welfare. As it avoids the Sadistic Conclusion, the Neutral-Range is still a better alternative to solve the Repugnant Conclusion than the Critical-Level Principle.

6 Conclusion

Within this dissertation I have shown that the solutions usually offered to solve the Repugnant Conclusion lead to even more counter-intuitive conclusions. This does not, however, mean that we must abandon our intuitions that lead us to the Repugnant Conclusion or accept it. By adapting the Critical-Level Principle and introducing a Neutral-Range I have created a solution to the Repugnant Conclusion which will avoid the Sadistic Conclusion, and respect our intuitions regarding egalitarianism, the moral significance of numbers and suffering. However, if we choose to avoid the Repugnant Conclusion using this principle we must accept that its application might be *ad hoc*. I have failed to offer an adequate account of why a Neutral-Range is not an arbitrary principle. I can only speculate that the application of the Neutral-Range will involve an appeal to pluralism.

References

- [1] Arrhenius, G. (2000) 'An Impossibility Theorem for Welfarist Axiology', in *Economics and Philosophy* 16: 247–266.
- [2] Blackorby, C., Bossert, W. & Donaldson, D. (2004) 'Overpopulation and the Critical-Level Approach', in T. Tännsjö & J. Ryberg (eds.), *The Repugnant Conclusion*. Netherlands: Kluwer Academic Publishers.
- [3] Broome, J. (2005) 'Should We Value Population', in *The Journal of Political Philosophy* 13(4): 399-413.
- [4] Parfit, D. (1984) *Reasons and Persons*. Oxford: Oxford University Press.

- [5] Parfit, D. (2004) 'Overpopulation and the Quality of Life', in T. Tännsjö & J. Ryberg (eds.), *The Repugnant Conclusion*. Netherlands: Kluwer Academic Publishers.

Socrates' Philosophical Mission in the *Apology*: Irony and Obedience

Corrado Musumeci
University of Liverpool

Introduction

When navigating the maze that is Socratic thought, one can too easily forget that Socrates is but human and not by nature divine,¹ for he did not spring “like a mushroom out of the earth”.² However, the ordinariness of Socrates – insofar as being a son, father and husband³ of modest upbringing – is ever eclipsed by his eccentric features that made him very much *atopos*.⁴ The most eccentric feature of all was his philosophical mission. It is clear that the peculiar characteristics of Socrates made him notably singular. He displayed exceptional continence, abstemiousness, physical robustness and inhuman endurance; he engaged in prolonged fits of rapt contemplation, and was somewhat of an oddity in appearance and for possession of his divine sign, the *daimonion*.⁵ However, it was his obedience to the divine command as the god-gifted midwife for the Athenians that was his most remarkable singularity. A task that would gain him infamy in Athens and fame today.

In describing the challenge of forming the truest biographical sketch of the historical Socrates, Godley observed – and I believe no more fitting descrip-

¹A momentary illusion for us but possibly lifelong for his admirers, particularly Alcibiades and Apollodorus. One may also be reminded of Horace Mann's aphorism that “To pity distress is but human; to relieve it is Godlike” ([15] p. 297). Of course, the irony being that Socrates caused distress (of the mind) in order to relieve it (from the soul).

²Hegel [10] p. 384.

³His parents were Sophroniscus and Phaenareteto, his three sons; Lamprocles, Sophroniscus and Menexenus, and his wife; Xanthippe (Diogenes [6]).

⁴See Kierkegaard for a humorously similar observation: “As is known, Socrates was a loafer who cared for neither world history nor astronomy . . . But he had plenty of time and enough eccentricity to be concerned about the merely human, a concern that, strangely enough, is considered an eccentricity among human beings, whereas it is not at all eccentric to be busy with world history, astronomy, and other such matters” (Kierkegaard [13] p. 83).

⁵Taylor [23].

tion has been given since – that to view Socrates through the scant works left to us is always to view him *eminus et quasi per transennas* (at a distance and as if through latticework).⁶ This obstacle is no less apparent in understanding Socratic divine obedience. Moreover, Socrates’ diametric nature and use of irony – to some even the embodiment of irony⁷ – leaves us further with a distorted image. The wisest Athenian who knew nothing, the most accomplished teacher who did not teach,⁸ a philosopher who did not write, the ugliest man with the most “gorgeous and wonderful”⁹ of thoughts, Athens’ gravest annoyance and their greatest saviour, he who “exists and yet again does not exist in world history, . . . the nothing from which the beginning must nevertheless begin”.¹⁰ In the Platonic dialogues, one is always distanced from Socrates by his ambiguity in sincerity and his mask of irony; of pretence, confiction and masquerade. A distance, it has been argued,¹¹ exacerbated in the *Apology*. Thus the study presented here attempts to steer through the endless webs of equivocation and paradoxical confusion in order to clarify the irony and obedience that characterised his philosophical mission.

I first look at the extent of irony in the Delphic oracle story – Socrates’ account of hearing that Apollo’s oracle in Delphi deemed no one to be wiser than him – and thus the difficulties in comprehending Socrates’ interpretation of the Pythia’s answer as a divine obligation, prioritised above even life itself. This close reading of Plato’s *Apology* will allow for a concluding portrait of Socrates and his philosophical mission. One that overwhelmed Athens and relegated all other aspects of Socrates’ life to afterthoughts, including his own family and political involvement, and that which ultimately led to poverty, hostility and state sanctioned execution.

Socrates has captured the minds of later thinkers and scholars in no dissimilar

⁶Godley [7].

⁷Most famously, Kierkegaard: “irony constituted the substance of his existence” ([12] p. 12).

⁸Nehamas ([17] p. 62) highlights an important distinction that while Socrates did not picture himself as a teacher (of *arête*) and denied to be attached to such a label, he was certainly “perceived as a teacher” or taken to be one indirectly, most obviously by Plato and Xenophon.

⁹Alcibiades compares Socrates to a figure of Silenus in appearance (with a snub nose and bulging eyes) but also that the figures have effigies of gods inside, in the same way Socrates’ insides are “divine, so glorious, so gorgeous and wonderful” (*Symposium*, 216d–217a).

¹⁰Kierkegaard [12] p. 198.

¹¹For example, Leibowitz [14].

way to his role as the stingray or the Marsyas¹² of Athens. In particular, the perplexing self-interpreted mission of a simple wise man, and its subsequent cause of estrangement and death in a democratic *polis*, has directed and sustained much interest. This study seeks to answer why Socrates interprets the oracle's answer as a divine command that ought to be prioritised above the requirements of the *polis*, the law, his family, his civic duty, his livelihood, and his own life in the face of death.

1 Irony in the Delphic oracle story

To understand Socrates' interpretation of the Delphic oracle story as a divine command of supreme magnitude, the question must first be settled of whether the story, and Socrates' treatment of it, is ironic or if earnestness has been mistaken for irony. Hackforth takes both the oracle story itself and Socrates' refutation and understanding of the oracle's answer as sincere and truthful.¹³ He does not believe Plato invented either, rather that Plato – mistakenly and groundlessly – attempted to use Apollo and the oracle's answer as the prescription and inspiration of Socrates' elenctic and protreptic activities, i.e. as the origin of his mission. For Hackforth, Socrates' "interpretation of the oracle is a typical example of his accustomed irony", for in testing the oracle, Socrates did not take seriously the answer and therefore did not "regard it as the voice of God".¹⁴

Reeve responds that it is possible for Socrates to take seriously the oracle while inquiring into its true meaning, pointing to the fact that Herodotus portrays the gods as welcoming the inspection of an oracle; indeed they are of the opinion that it is wise to do so.¹⁵ Hence, it is not impious but also not ironical. Moreover, after asserting that Socrates did not regard the answer as the voice of God, Hackforth goes on to say, "his obedience to the voice of God

¹²Those who hear the words of Socrates are "overwhelmed and spellbound" in a way comparable to the pipe playing of Marsyas (*Symposium*, 215c–d). However, it is interesting to speculate as to whether Plato was perhaps also being ironic. Socrates – the man who by the divine command of Apollo endeavoured on an anti-hubristic mission to his death – is being compared to Marsyas – who in his hubris, Apollodorus (I.4, 2) tells us, challenged the god Apollo to a musical contest; only to be defeated and suffer death.

¹³Hackforth [9].

¹⁴Ibid. p. 94.

¹⁵Reeve [22] p. 23.

... contained no element of irony".¹⁶ How can his obedience to something which he does not regard seriously, contain no irony? Socrates cannot be ironic in how he regards the answer and then not be ironic when adhering to the answer. If Hackforth believes his obedience was not ironic then this presupposes, contrary to his initial claim, that his interpretation of the oracle was not ironic.

However, there are others who support Hackforth.¹⁷ Taylor believes the apology to only begin seriously at 28a, before which there is an abundance of that "whimsical humour" of Socrates' irony, wherein he is "not wholly serious".¹⁸ The oracle story in particular is relegated to "native humour".¹⁹ Taylor's argument for this relies on the distinction in Socrates' choice of wording. Before 28a, Socrates speaks of the command received by "the gods" and "Apollo" but from 28a – that is, in what Taylor calls the actual or serious defence – by "God". Taylor's translation must differ from that of Tredennick & Tarrant used here, for at 23c2 Socrates speaks of his "service to God". Ignoring this, it is still evident that Taylor neglects the fact that Socrates speaks of the oracle before and after 28a frequently, which is no different to speaking of Apollo. Lastly, Taylor does not allow for Socrates to be both truthful and humorous at once. West asserts that "Socrates' speech in the *Apology* is both ironic and boastful, and he jokes while telling the truth".²⁰ This observation is most apparent when Socrates boastfully compares himself to Heracles at 22a8; it is arrogant and humorous yet need not require us to question the truth of his onerous task.

The view that Socrates was speaking ironically here suggests that Plato has depicted a man who continually refers to the truth as a matter of considerable importance yet speaks falsely in his own defence. Socrates is unceasingly adamant from the very start of the *Apology* that "scarcely a word of what [his accusers] said was true" (17a4); that they "have said little or nothing that is true" (17b8-9); that the accusers are "without a word of truth" (18b4) and spout "untrue accusations against [him]" (18b9). All the while, Socrates assures the jury that from him "you shall hear the whole truth" (17b9) with no "flowery language" (17b10) and further, that he has the highest confidence "in

¹⁶Hackforth [9].

¹⁷See also I.G. Kidd and John Burnet.

¹⁸Taylor [24] p. 160.

¹⁹Ibid.

²⁰West [27] p. 104.

the justice of [his] cause” (17c3). He later repeats his sincerity for emphasis at 20d4–6 just before telling the story itself: “Perhaps some of you will think that I am not being serious; but I assure you that I am going to tell you the whole truth”. First, it is clear that if this is ironic it is certainly not the characteristic Socratic irony we are accustomed to for it is not remotely subtle or humorous but a forthright lie. Second, if we concede that the Delphic oracle story or Socrates’ subsequent treatment of it are by nature ironic, then we must agree to Plato depicting a hypocritical and contradictory Socrates. One who cannot value truth for he outright deceives the jury while declaring the value he places on truth and justice:

[A man] ought to inform [the jury] of the facts and convince them by argument. The jury does not sit to dispense justice as a favour, but to decide where justice lies [...] It follows that we must not develop in you, nor you allow to grow in yourselves, the habit of perjury; that would be impious for us both. (35c2–9)

Moreover, it would portray Plato not as a masterly writer vindicating his wise teacher but one who has mistakenly created a Socrates far from the figure of an ideal sage; no more truthful or just than those he previously denounced. Brickhouse & Smith rightly argue that “it is most unlikely that he would resort to mere irony on a matter of such substance to his defence”.²¹ It would be nonsensical for Socrates to advocate truth and justice while being (pointlessly) deceptive on a topic so crucial to his defence. It seems the only real irony is in the charges he faces. His actions are being accused as impious yet these very actions are his efforts to maintain piety. To be insincere and display mockery of the gods, in a courtroom no less, would be impiety that Socrates maintains to avoid (for example at 35d3–18).²²

2 From Descriptive to Prescriptive

Grasping why Socrates translated the oracle’s answer as an ultimately important divine obligation is essential to understanding Socrates’ philosophical life as the self-acknowledged gadfly, “which a god had placed on the neck of the

²¹Brickhouse & Smith [4] p. 658.

²²“Surely it is obvious that if I tried to persuade you and prevail upon you by my entreaties to go against your solemn oath, I should be teaching you contempt for religion; and by my very defence I should be virtually accusing myself of having no religious belief. But that is very far from the truth. I have a more sincere belief, gentlemen, than any of my accusers.” (35d3–18).

beautiful steed Athens that it might never be allowed any rest”.²³ The oracle certainly does not appear prescriptive and amounts to a short descriptive answer that “there was no one” wiser than Socrates (21a8). This has aroused reasonable confusion for Socrates has made quite a speculative leap from hearing that no one is wiser than himself, to finding it his lifelong divine obligation to philosophise – to rid the Athenians from their ignorance of lacking “more than human” wisdom, and thus exhort them to tend to their souls through understanding the truth of the good and noble.

Indeed, Vlastos compares it to “pulling a rabbit out of a hat”²⁴ and Hackforth²⁵ feels “it is by no means clear why Socrates should speak of his examination of the reputed wise men as obedience to the god”.²⁶ Socrates repeats this notion of obedience to and obeying of God again and again throughout his speech, totalling no less than on sixteen occasions (22a4, 23b5–6, 23b8, 23b10, 28d9–11, 28e4–7, 29a2, 29b8–9, 29d2–6, 30a6–8, 30e3–7, 31a7–8, 31a10, 33c5–7, 35c11, 37e7–8). The only matter Socrates returns to as frequently is his proclaimed ignorance; accordingly to ignore that which concerned him so greatly in his final moments of freedom would surely be an error. Nehamas’ examination of how Socrates was *able* to come to this realisation that the god commands him will be the starting point.²⁷ From there, the question of *why* Socrates interprets the god as he does will be resolved.

2.1 Realising the Command

In an essay that aroused “friendly though spirited” criticism from Vlastos, Nehamas²⁸ argues that Socrates was able to heed the god’s good wish to tend to his imperfect soul by virtue of luck. For Nehamas, there was in fact no initial reason at all why “Socrates was correctly motivated to inquire into virtue so as to be able to interpret the god’s command correctly”.²⁹ If this is true, it is on the face of it a somewhat underwhelming conclusion – that so simple an

²³Nietzsche [18] p. 159.

²⁴Vlastos [25] p. 171.

²⁵Hackforth [9] p. 89.

²⁶Ibid. p. 89.

²⁷Nehamas [17].

²⁸Ibid.

²⁹Ibid. p. 77.

answer can be given to so complex an arrangement – but that alone is no reason to discredit it. Vlastos’ solution is to the contrary of Nehamas, wherein he reasons that, for the god to successfully transmit his desire, he depends on someone who can bring “the right beliefs to the interpretation of [divine] signs” in order to read them correctly and assist the god with what he wills for Athenians. That which he “would be doing himself in person, if he only could”.³⁰ Vlastos is asserting that Socrates, with his right beliefs, was able to interpret the oracle in order to read its correct meaning. In doing so, Socrates envisaged a divine responsibility. To this, Nehamas poses the logical question of how Socrates acquired these true beliefs in the first instance – tentatively³¹ clarifying the answer as from the result of “street-philosophising”.³² Nehamas then goes on to criticize Vlastos by arguing that this “does not explain why Socrates engaged in such “street-philosophising” in the first place”.³³ But this is not a valid response for two reasons: it fails to undermine the argument at hand and, regardless of this, it can quite readily be explained.

There are a handful of reasons Socrates may have engaged in street-philosophising (that is, the search for ethical truth) prior to receipt of the Delphic oracle. It may have simply been the inevitable fruit of combining Socrates’ inquisitive nature with the intellectual environment of Athens; that “metropolis of wisdom”;³⁴ that city “which is the greatest and most famous in the world for

³⁰Vlastos [25] p. 173–4.

³¹I use ‘tentatively’ because Nehamas says that Vlastos “seems to suggest” (Nehamas [17] p. 82) as if he is not wholly certain of what Vlastos is in actuality suggesting. Vlastos does not in fact suggest that but it is a reasonable presumption that Socrates engaged in street-philosophising before the oracle’s answer. One must therefore wonder if Nehamas has misinterpreted the context and time of the Socrates that Vlastos attributes the term “street-philosophising” to. Vlastos has in mind Socrates’ activity of questioning the politicians, poets and craftsmen after hearing the oracle’s answer and thus using this experience to interpret the divine task. On the other hand, Nehamas is using the term “street-philosophising” to explain Socrates’ acquiring of true beliefs before hearing the oracle’s answer. While this is a reasonable extrapolation to make, it is not strictly in line with the sequence of argument that Vlastos explicates (as Vlastos himself does not attempt to answer where these true beliefs arose). In any case, we know that Socrates was considered wise prior to Chaerephon’s question at Delphi and that Socrates had searched for moral truth “else he could not have read correctly the signs the god sent him.” (Vlastos [25] p. 173) Consequently, the formation of true beliefs was inevitable.

³²Nehamas [17].

³³Ibid.

³⁴*Protagoras*, 337d.

its wisdom”;³⁵ that place “of resort for the thinkers of the Hellenic world”;³⁶ that city which “revealed philosophy” and “honoured the power of discourse” whereby “people are called Greeks because they share in our education rather than in our birth”.³⁷ It may have been because, as Aristotle reminds us, “By nature, all men long to know”³⁸ – and what better subject to desire knowledge of than goodness itself. It was quite possibly because Socrates was encouraged, indirectly or directly, by those great minds before him; Parmenides, Zeno, Pericles, Anaxagoras, Archelaus.³⁹ Speculating an answer to this question is, however, unnecessary because the root of Socrates’ quest for ethical fundamental truths is insignificant here. What *is* significant is that it created right beliefs which, in turn, made it “possible for him to guess that a unique responsibility was laid on him to use in the god’s service what little wisdom he has”.⁴⁰ We ought, then, to describe as lucky not that which Nehamas propounds but rather that which Nehamas criticises Vlastos over. Namely, Socrates’ early motivation to inquire into virtue before receipt of the Delphic oracle – just as he did many topics of a fundamental and puzzling nature and did so at the forefront of Athenian intellectualism – resulting in possession of true beliefs.

2.2 Understanding Socrates’ Interpretation

Why does Socrates interpret the oracle as a command requiring unfaltering obedience? An answer can seemingly be found via reconciling two seminal studies. Brickhouse & Smith argue that Socrates’ references to obeying a divine command display the oracle as the “seminal element in the genesis of Socrates’ understanding of his specific mission”⁴¹ and that “his certainty about the moral importance of his mission is derived from various forms of divina-

³⁵ *Apology*, 29d8–9.

³⁶ Taylor [23] p. 51.

³⁷ Isocrates [11] *Ad Demonicum*, pp. 47–50.

³⁸ *The Metaphysics*, 980a.

³⁹ For a quite brilliant biography of Socrates’ early life, the reader need not look further than Taylor [23]. It is pertinent here to emphasise Taylor’s inference that “the fame of Socrates as a man of outstanding intellectual power must have been established in this early half of his life” (p. 72). Nehamas’ question of why exactly Socrates engaged in street-philosophising would presumably be self-evident to Taylor.

⁴⁰ Vlastos [25] p. 174.

⁴¹ Brickhouse & Smith [3] p. 88.

tion”.⁴² Reeve initially identifies this explanation as based in religion rather than in Socrates’ conception of virtue derived from the *elenchus*.⁴³ However, he oversteps this when suggesting that Brickhouse & Smith believe Socrates to perform his mission “simply and solely because Apollo has ordered him to”.⁴⁴ In fact, they directly contradict such an interpretation: “Socrates did not see the oracle as a divine command issued to him by the god”.⁴⁵ Instead, Socrates had prior beliefs of piety as requiring “that a man engage in just service to the god” or requiring “one to serve the gods by promoting what is good.” Once Socrates had reasoned the god’s desire for man, his belief of piety required him to serve the god by freeing Athenians “from the pretence of real wisdom [and urging] them to acquire virtue”. In this way, Socrates is carrying out “his pious duty to convey the god’s message”.⁴⁶ Of course, there is a fine (but distinct) line between (a) Apollo ordering Socrates to conduct his mission and (b) Socrates, in the interest of maintaining piety – where piety requires serving the gods via promoting goodness – acting as a servant to the god with the intention of helping fulfil what Socrates believed “the god wanted most for men”.⁴⁷

Reeve’s explanation is seemingly founded on ethical reasons first. Virtue not religion is the basis of Socrates’ mission because, as Reeve infers from the *Euthyphro*, “virtues are not just independent of the gods [who themselves are good and virtuous], they are more important”.⁴⁸ Further, the sole deciding factor in performing any action is whether one is acting justly or unjustly (28b8–c2). Thus, these conclusions arrived at via the *elenchus* provide the crux of Reeve’s proposal, which is worth quoting at length:

The gods are good and virtuous, [therefore] whatever they command must itself be good and virtuous. The demands of virtue override all others, [therefore] we must obey those commands. But because virtue is independent of the approval or disapproval of the gods, the authority divine commands have over us derives

⁴²Ibid. p. 107.

⁴³Reeve [22].

⁴⁴Brickhouse & Smith [3] p. 62.

⁴⁵Ibid. p. 99.

⁴⁶Ibid. pp. 97–98.

⁴⁷Ibid. p. 99.

⁴⁸Reeve [22] p. 66.

entirely from the independent and overriding ethical authority of the virtues [...] We should obey them because, on the basis of elenctic argument, [...] we have an independent ethical reason to do so.

And, as Socrates expresses that “to do wrong and to disobey my superior, whether god or man, is bad and dishonourable” (29b8–9), it follows that he will not do so. However, the distinction between the explanations from Reeve and Brickhouse & Smith has now become less pronounced. They are unanimous that Socrates’ mission began with (and continued because of) the answer from the Pythian priestess. Here, Reeve sees Socrates adhering to ethical reasoning behind the divine command. Brickhouse & Smith see Socrates adhering to the pious obligation of conveying the god’s message – which is entirely virtuous. Hence, the apparent fundamental distinction lies in Brickhouse & Smith allegedly rejecting that which Reeve concentrates on; namely, that Socrates’ duty derives from his conception of virtue from the *elenchus* (as opposed to from divination). However, we can see a reconciliation of the two views emerge and, in a sense, Reeve’s explanation subsuming that from Brickhouse & Smith. The latter actually support the former yet conclude to the contrary. This can be seen in the detailing of their view. Brickhouse & Smith assert two essential and very telling points:⁴⁹

1. The god’s message ... is that the Athenians lack the supreme benefit, virtue, a deficiency of which they are arrogantly ignorant.
2. As the god’s servant, Socrates must free them from the pretense of real wisdom [...] and urge them to acquire virtue.

Piety “requires one to attempt to aid the gods by carrying out what the gods ordain”⁵⁰ and the gods ordain, as Reeve shows, only the good and virtuous. The conclusion being that Socrates has carried out this pious service to convey and aid the god’s virtuous message. Do we not but see Reeve’s explanation here? – virtue as the overriding authority and, derived from this, the authority of the god’s divine command. Now one may argue that the Socrates of Brickhouse & Smith does not hold that what the gods ordain must be virtuous. What is quite clear in the *Apology*, however, is that Socrates does believe this specific message to be virtuous, regardless of whether he believes *all* commands to be virtuous. If, as Socrates asserts at 28b8–c2, the sole determining

⁴⁹Brickhouse & Smith [3] p. 98.

⁵⁰Ibid. p. 99.

factor for an action is whether one is acting justly or unjustly, then Socrates must hold that the god's message is a virtuous one, for otherwise his actions – that have promoted this message – would be unjust. Socrates defends these actions relentlessly; describing himself (indirectly) as “the true champion of justice” (32a). Indeed, justice for Socrates is paramount and drives his defence speech until the close:

For Socrates no sacrifice is too great when justice or philosophy is at stake: he will neither do injustice nor abandon the life of philosophy though the pursuit of justice and philosophy cost him his life. Socrates' unwavering commitment to justice and philosophy determines his thinking on every important human question.⁵¹

Due to his sheer confidence “in the justice of [his] cause” (17c3), it is Socrates himself who reminds the jury that they “decide where justice lies . . . to return a just and lawful verdict” (35c4–6). Thus it follows that Socrates' fulfilment of the god's command was, from his perspective, just and the command itself a virtuous one. Obeying the god's command is “to engage in what he believes independently to be the greatest good for a human being: philosophy”.⁵² It is this picture of Socrates that is most fitting; a religious man who holds philosophy not *above* faith as Reeve concludes, but juxtaposed with faith. The great street-dialectician who, in order to lead and urge the life of virtue and truth, will follow his *daimonion* and divine command unquestionably and no less than he will follow the persuasion of argument, reason and reflection.⁵³ Thus, Socrates is not willing to cease practising philosophy because first; it is to disobey the god (37e6), and second; “this even happens to be the greatest good for a human being”.⁵⁴

This is in agreement with Vlastos who sees sense and coherence in Socrates' equal unconditional commitment to both obeying divine commands and fol-

⁵¹Weiss [26] p. 2.

⁵²Ibid. p. 16.

⁵³See *Crito* (46b4-7): “my attitude is not unprecedented, for it's always been my nature never to accept advice from any of my 'friends' except the argument that seems best on reflection”. A not entirely relevant but certainly interesting anecdote is told by Cicero ([5], 4.80), which also emphasises the importance and power of reason for Socrates: “A certain Zopyrus claimed that he could discern a person's nature from his physiognomy. This man gave out a list of Socrates' faults in the midst of a gathering and was laughed at by all the rest, for they were aware that Socrates did not exhibit those faults. Socrates himself, however, supported Zopyrus, saying that they were indeed inborn in him, but that he had cast them out by reason”.

⁵⁴Weiss [26] p. 16.

lowing reasoned argument: “these two commitments cannot conflict because only by the use of his own critical reason can Socrates determine the true meaning of any of these signs”.⁵⁵ Just as our solution depicts Socrates using *elenchus*-based ethical reasons for obeying the oracle, Vlastos understands that the process by which Socrates interprets the real meaning of the oracle – that is, as a divine command – is thoroughly subjective, and points to 28e4 for evidence: “God appointed me, as I *supposed* and *assumed*, to the duty of leading the philosophic life, examining myself and others”. Further, this was one of many divine signs: “This duty I have accepted, as I said, in obedience to God’s commands given in oracles and dreams” (33c4–7). Vlastos identifies that even if we assume that the direct command which is absent in the *Apology* was clearly spelled out in a dream, it would still require reason to give him certainty.⁵⁶ The command is still dependent on reason. Reeve believes Brickhouse & Smith to have “succumbed to [the] temptation”⁵⁷ of concluding Socrates’ mission as wholly religious. This is not the case as we have shown and Vlastos confirms it. Socrates is religious insofar as he upholds the religious expectations he was born into as a citizen of Athens and has various divine encounters or conversations. However, it is critical reason that allowed the command to reach Socrates, not through a direct command from Apollo. It was Socrates’ implacable drive via piety and reasoned argument toward goodness and truth that not only found his divine command, but sustained it through resistance until death.

However, Socrates is not “counting on two disparate avenues of knowledge about the gods”.⁵⁸ He does not adhere “uncompromisingly to the authority of reason”⁵⁹ in the first instance. Rather, the ambiguous divine signs (including the Delphic oracle) are “susceptible [to] alternative interpretations”,⁶⁰ specified and chosen by the authority of *elenchus*-based ethical reasons that Reeve also asserts. Before closing, Vlastos asks: “Would Socrates have given his life to this task if his piety had not driven him to it?”⁶¹ The answer, we must conclude, is no – in the same way that Socrates would not have given his life to

⁵⁵Vlastos [25] p. 171.

⁵⁶ Vlastos [25] p. 172.

⁵⁷Brickhouse & Smith [3] p. 62.

⁵⁸Ibid. p. 167.

⁵⁹Ibid. p. 166.

⁶⁰Ibid. p. 168.

⁶¹Ibid. p. 177.

his task if reason had not driven him to it. Socrates' mission cannot have piety without reason; divine command without the *elenchus*; religion without *logos*. While the divine signs prompted Socrates, it was the persuasion of the overriding authority of reasoned argument and virtue that equally (and ultimately) decided his philosophical life and subsequent death.

The strongest textual evidence to support this conclusion to Socrates' divine command is at 28d7–11, where Socrates, in illustrating his implacable commitment to the mission, defiantly affirms:

Where a man has once taken up his stand, either because it seems best to him or in obedience to his orders, there I believe he is bound to remain and face the danger, taking no account of death or anything else before dishonour.

Here, “obedience to his orders” is not reference to a command from Apollo to Socrates, but rather is meant as *military* orders. This is made clear for in the same section Socrates is comparing his defiance in the face of death to that of Achilles and other “heroes who died at Troy” (28c3). Thus at 28d7 Socrates is simply concluding his point, which he does so even more emphatically at 30b9–c2: “I am not going to alter my conduct, not even if I have to die a hundred deaths”. Therefore, Socrates is only displaying two reasons for a man to be in his position (that is, a man who has “taken up his stand”); either it is from reason or from military orders. It cannot be the latter and so must be from reason, which directly supports the solution of Reeve *et al.*⁶²

3 Conclusion

Socrates' conception of his mission rests on two pillars. The first is divine communication through oracles and dreams and, independently but overriding this, critical reason. Socrates obeys the god's command but has subjectively interpreted it through the authority of reasoned argument and ethical reasons. Therefore, we must agree with Weiss that to construe Apollo as directly commanding a servile Socrates is a misreading of the Delphic oracle story and its subsequent treatment.⁶³ Apollo “does not serve as an alternative to human reason as a guide to moral decision making; instead, the god issues a divine

⁶²“A man who, as he says, ‘is persuaded by nothing in me except the argument that when I reason seems to me to be the best’ (Crito, 46b4–6).” (Reeve [22] p. 73).

⁶³As well as Reeve [22] p. 62, Brickhouse & Smith [3] p. 99, and Vlastos [25].

mandate to use human reason as one's guide".⁶⁴ It follows that the figure of Socrates and his philosophical endeavours are to be viewed as much based in religious duty as based in ethical reason, though it is the independent authority of the latter that specifies and overrides the former.

It is because Socrates maintained so uncompromisingly his adherence to goodness and justice through piety and truthful examination that his *daimonion* remained utterly silent after the guilty verdict.⁶⁵ Socrates died not with his mission complete but with it honorably sustained, and with his admissions of ignorance and obedience consistent with his prioritisation of reasoned argument, virtue, justice, piety and truth. It is difficult to imagine that what awaited Socrates was the prison of Tartarus, not the Isles of the Blest. It is found then that not only is it "better to be Socrates dissatisfied than a fool satisfied"⁶⁶ but that this dissatisfaction in life brought about satisfaction at death. That his mission was a just one, as he began the defence proclaiming (17c3), is confirmed once more at the speech's close not only by his divine sign failing to oppose him but that he does not succumb to using pitiful tactics of obtaining acquittal, nor does the prospect of death fill him with regret. For these are the reactions only of a guilty man. Like the swans who belong to Apollo, Socrates is "dedicated to the same god" and is "no more disconsolate at leaving this life" than them.⁶⁷ The unexamined life is not worth living⁶⁸ but equally; the unexamined life is not worth dying for.

References

- [1] Apollodorus. (2008) *The Library of Greek Mythology*, R. Hard (trans.). Oxford: Oxford University Press.
- [2] Aristotle. (2004) *The Metaphysics*, H. Lawson-Tancred (trans.). Penguin Books.
- [3] Brickhouse, T.C. and Smith, N.D. (1989) *Socrates on Trial*. Oxford: Clarendon.

⁶⁴Weiss [26] p. 16.

⁶⁵See 40a7–12.

⁶⁶Mill [16] p. 281.

⁶⁷Plato [19] *Phaedo*, 85b4–8.

⁶⁸*Ibid.* 38a.

- [4] Brickhouse, T.C. and Smith, N.D. (1983) 'The Origin of Socrates' Mission', in *Journal of the History of Ideas* 44(4): 657-666.
- [5] Cicero. (2002) *Cicero On The Emotions: Tusculan Disputations 3 and 4*, M. Graver (trans.). Chicago, IL: University of Chicago Press.
- [6] Diogenes, L. (2011) *The Lives and Opinions of Eminent Philosophers: Volume II*, R.D. Hicks (trans.). Witch Books.
- [7] Godley, A.D. (1896) *Socrates And Athenian Society In His Day: A Biographical Sketch*. London: Seeley.
- [8] Grote, G. (1859) *The Life, Teachings, and Death of Socrates from Grote's History of Greece*. New York, NY: Stanford & Delisser.
- [9] Hackforth, R. (1933) *The Composition of Plato's Apology*. Cambridge: Cambridge University Press.
- [10] Hegel, G.W.F. (1995) *Lectures on the History of Philosophy: Greek Philosophy to Plato*, E.S. Haldane (trans.). Nebraska: University of Nebraska Press.
- [11] Isocrates. (2004) *Isocrates I & II*, D. Mirhady, Y. Too & T. Papillon (trans.). Austin, TX: University of Texas Press.
- [12] Kierkegaard, S. (1992a) *The Concept of Irony with Continual Reference to Socrates*, H.V. Hong and E.H. Hong (trans.). Princeton, NJ: Princeton University Press.
- [13] Kierkegaard, S. (1992b) *Concluding Unscientific Postscript to Philosophical Fragments*, H.V. Hong and E.H. Hong (trans.). Princeton, NJ: Princeton University Press.
- [14] Leibowitz, D.M. (2010) *The Ironic Defence of Socrates: Plato's Apology*. Cambridge: Cambridge University Press.
- [15] Mann, H. (2012) *Lectures on Education*. Reprint, Ulan Press.
- [16] Mill, J.S. (2004) *Utilitarianism and Other Essays*, A. Ryan (ed.). Harmondsworth: Penguin Books.
- [17] Nehamas, A. (1999) *Virtues of Authenticity: Essays on Plato and Socrates*. Princeton, NJ: Princeton University Press.
- [18] Nietzsche, F. (1996) *Human, All Too Human: A Book for Free Spirits*, R.J. Hollingdale (trans.). Cambridge: Cambridge University Press.
- [19] Plato (2002) *Five Dialogues: Euthyphro, Apology, Crito, Meno, Phaedo*, G.M.A. Grube (trans.). Indianapolis, IN: Hackett.

- [20] Plato (2003) *The Last Days of Socrates*, H. Tredennick and H. Tarrant (trans.). London: Penguin Books.
- [21] Plato (2008) *Symposium*, R. Waterfield (trans.). Oxford: Oxford University Press.
- [22] Reeve, C.D.C. (1989) *Socrates in the Apology: An Essay on Plato's Apology of Socrates*. Indianapolis, IN: Hackett.
- [23] Taylor, A.E. (1932) *Socrates*. Edinburgh: Peter Davies.
- [24] Taylor, A.E. (1937) *Plato: The Man And His Work*. London: Methuen.
- [25] Vlastos, G. (1991) *Socrates: Ironist and Moral Philosopher*. Cambridge: Cambridge University Press.
- [26] Weiss, R. (1998) *Socrates Dissatisfied: An Analysis of Plato's Crito*. Oxford: Oxford University Press.
- [27] West, T.G. (1979) *Plato's Apology of Socrates*. Ithaca, NY: Cornell University Press.

Environmental Philosophy and the Teleological Argument: A Polemic*†

Thomas E. Randall
University of Lancaster

1 Introduction

Within this paper, I argue that the teleological argument, understood within an environmental philosophical context, is a weak theoretical attempt to attribute intrinsic value to biological organisms, and holistic environmental entities. As the teleological argument characterises the fundamental premise of biocentrism (in both its individualistic and pluralistic approaches) and ecocentrism, my following critique draws attention to the legitimacy of these two perspectives in being serious contenders as coherent environmental ethical theories. For if my following argument is accepted, this ultimately undermines the aforementioned philosophical accounts on how the environment ought to be treated, or even perceived as being.

In arguing my point, I shall begin by outlining the teleological argument and its importance to the theoretical underpinning of biocentrism and ecocentrism, making distinct the argument's differential use quite clear. Subsequently, I attack the teleological argument from two angles: the first focuses on the idea that interpreting an organism or holistic environmental entity in a teleological manner raises awkward conclusions that demonstrate the absurdity of its usage; the second regards the erroneous use of metaphoric language within the above two environmental ethical theories that constructs the perceived utility of the teleological argument within their approach.

*Delivered at the BUPS Autumn Conference 2013 on 23–24 November at the University of Manchester.

†I sincerely thank Josh Milburn (PhD student, Queens University, Belfast) for his time spent giving extremely helpful constructive feedback.

2 Environmental Ethics and Intrinsic Value

To identify how the teleological argument forms the basis of these two environmental ethical theories, it is first important to elucidate the disparity and the commonalities that exist between biocentrism and ecocentrism to exemplify this. Indeed, what draws these two theories away from each other, yet what leads both to make use of the teleological argument, is the interpretation of which environmental entity possesses intrinsic value.

Biocentrism, simply defined, is an ethical theory that attributes intrinsic value to all biological organisms.¹ As stated in the introduction, biocentrism splits into two main factions: biocentric individualism and biocentric pluralism, with this partition purely centred upon which environmental entities are to be perceived as being intrinsically valuable. The individualist account argues in favour of all individual biological organisms possessing intrinsic value, but not the sum of their collective group. As Paul Taylor writes within *Respect for Nature*:

The population has no good of its own, independently of the good of its members. To promote or protect the population's good ... does not mean that the good of every one of its members is also promoted and protected.²

Interestingly, biocentric pluralism develops as a response to Taylor's sole focus on the intrinsic value of the individual organism. James Sterba provides a critique of Taylor in this manner, stating that "species can be benefited and harmed and have a good of their own, and so qualify on Taylor's view as moral subjects".³ As such, the biocentric pluralist emphasises the intrinsic worth of larger environmental entities, as compared to an individualist account that ascribes none whatsoever.

Ecocentrism, though closely inferred as relating to biocentric pluralism, "goes beyond biocentrism", by which is meant, according to Stan Rowe, that:

The whole ecosphere is even more significant and consequential: more inclusive, more complex, more integrated, more creative,

¹Taylor [18] p. 12.

²Ibid. p. 69.

³Sterba [15] p. 192.

more beautiful, more mysterious, and older than time.⁴

Thus, ecocentrism provides an almost spiritual account as to how the environment ought to be perceived, to be contrasted with a far more scientifically grounded biocentric theory.⁵ The entire ecosystem is to be attributed intrinsic value in the maintenance and stability of its existence, ahead of smaller considered environmental entities.

Despite this division, both theories originate from the same antagonistic position against what can be seen as being the prevailing environmental ethical discourse of anthropocentrism. Such unease from the biocentric and ecocentric outlook of anthropocentric thought develops out of the joint belief that this dominant environmental ethic exhibits an inadequate approach in understanding how humans ought to interact with the natural environment. This concern regards the anthropocentric perspective that only humans are to be attributed intrinsic value; consequently, nonhuman biological organisms, and holistic environmental entities, have solely instrumental worth.⁶ Moral consideration is granted only towards the beneficiary of human beings, and to none other.

Such an approach, it is argued by biocentrists and ecocentrists, has been the root cause of the vast number of environmental issues throughout human history. Both these two environmental ethical theories are apprehensive of the exploitation and destruction of the natural environment that the anthropocentric perspective has seemingly justified; “that man has been granted dominion over nature, that the plants, animals and ecosystems exist for us to *use*” [Scruton’s italics].⁷ As Taylor clearly delineates, “we will see the natural environment of our planet turned into a vast artifact”⁸ if humans are to continue to perceive the environment as “nothing more than resources for our own use and consumption”.⁹ In a comparable manner, ecocentrists equally emphasise the interconnectedness of nature, with deep ecologists Bill Devall and George Sessions putting it simply: “if we harm the rest of nature, then we are harm-

⁴Rowe [13] p. 106.

⁵Taylor [18] p. 111, in particular Taylor’s references towards closely linking biocentrism and the biological sciences.

⁶Norton [11] p. 164.

⁷Scruton [14] p. 197.

⁸Taylor [18] p. 5.

⁹Ibid. p. 12.

ing ourselves”.¹⁰ Furthermore, as David Worster highlights, the ecocentrist is perhaps seen in greater stead of challenging the noticeably arrogant perspective that anthropocentrism encourages of humans, aiming to counteract “this domination ideology”¹¹ through an upheaval of the anthropocentric mindset. These arguments have increasingly held greater sway within contemporary society, in the push for a more environmentally friendly global culture and greater ecological awareness.

This above similarity of biocentrism and ecocentrism in countering anthropocentric thought is important to accentuate, for it is this point that leads to these theories’ deliberation of extending the possession of intrinsic value beyond humans. This is done through the use of the teleological argument, which I will next proceed to explicate; in particular, I shall underscore how critical this argument is to both environmental ethical theories in substantiating their position on the attribution of intrinsic value, and thus how its subversion will notably knock unsteady the foundation of biocentrism and ecocentrism.

3 The Teleological Argument

Biocentric individualism displays the most obvious use of the teleological argument as its basic foundation; as Taylor writes, this argument is “central to the biocentric outlook”.¹² Hence, it will be sensible to explore this exemplar first of all as to highlight more evidently the occurrence of this argument within biocentric pluralism and ecocentrism. Within an environmental philosophical context, the teleological argument is to be understood in the Aristotelian sense; all biological organisms “seen to be a teleological (goal-orientated) centre of life”¹³ are conceived as “striving to preserve [themselves] and realise [their] good in [their] own unique way”.¹⁴ It is the flourishing of biological, and where appropriate psychological, functions of all living organisms, be they human, nonhuman animal, or plant. This way of viewing organisms as having a form of *telos* does not entail anthropomorphising, as Taylor is intent to

¹⁰Devall & Sessions [6] p. 68.

¹¹Worster [22] p. 44.

¹²Taylor [17] p. 237.

¹³Taylor [18] p. 45.

¹⁴Ibid. p. 121.

demonstrate. This teleological function of all organisms will continue to exist, whether or not human perception of this phenomenon is present.

It is this above argument that forms the first premise for attributing intrinsic value to nonhuman organisms. In understanding that all organisms possess a *telos*, the biocentric individualist asserts that any negative interference committed upon a living organism in the pursuit of their individual good induces a harm – and this is regardless of the capacity to experience physical or psychological pleasure or pain. Kenneth Goodpaster, in outlining this approach, affirms that to recognise all organisms as teleological centres of life is to thus recognise that all organisms have “capacities for benefit and harm”¹⁵ regarding any enhancement or hindrance, respectively, to the realisation of their good. It is this interpretation of harm given that invokes moral consideration to all organisms.¹⁶ On this view, intrinsic value exists individually within all teleologically conceived organisms; each is to be morally considered in virtue of the pursuit of their own individual good, described by Taylor as one having a respect for nature.¹⁷

Biocentric pluralism, however, avers that the logic of the individualist argument extends beyond the limits of the above analysis. Sterba argues that much larger environmental entities, such as species, possess their own good independent of the individual components that it is comprised of; species “evolve, split, bud off new species, become endangered, go extinct, and have interests distinct from the interests of their members”.¹⁸ While the last individual organism of a species can suffer pain at death, this is a separate interest than one that can be attributed to the species’ interest as a whole: that of not becoming extinct. It is therefore clear that in purposefully endangering a species, this harms its *telos* of flourishing and surviving. In this way, the biocentric pluralist puts forward that it is not just individual organisms that can be seen as teleological.

Broadening the scope of intrinsic value even further, the teleological argument is used by ecocentrists in reference to whole ecosystems. While biocentric pluralism believes the individualist argument is cut too short, ecocentrism seeks to rectify what is seen as an inadequacy of the individualist argument in fail-

¹⁵Goodpaster [8] p. 320.

¹⁶Varner [20] p. 113.

¹⁷Taylor [18] pp. 44–46.

¹⁸Sterba [15] p. 192.

ing to accommodate conservation concerns for ecological wholes.¹⁹ While slightly less referenced than within biocentrism, the teleological argument becomes apparent when we consider the following widely-quoted statement by Aldo Leopold's development of a land ethic: "A thing is right when it tends to preserve the integrity, stability, and beauty of the biotic community. It is wrong when it tends otherwise".²⁰ The *telos* of an ecosystem relates to reaching a particular equilibrium of stability; to divert this goal is to cause harm upon this entity. Individualism is too reductionist an approach to aid the eco-centric objective of whole ecosystems achieving their good, in maintaining a constant flourishing composition.²¹

4 Critiquing the Teleological Argument

Now it has been shown how the teleological argument forms the premise for determining why it is biocentrism and ecocentrism argue for the existence of intrinsic value within individual organisms and holistic environmental entities, I shall now seek to demonstrate why this is not a strong line of reasoning. It is my purpose, in the remainder of this paper, to convince the reader that the teleological argument is a weak foundation for both environmental ethical theories within the two points I before described: that (1) interpreting an organism or holistic environmental entity in a teleological manner leads to absurd conclusions, and (2) the use of metaphoric language within biocentrism and ecocentrism, in using the teleological argument, is unsound in its attempt of appeal for intrinsic value to be granted.

Beginning with biocentric individualism, to identify why the logic of the teleological argument forms asinine outcomes one must consider the wider implications of interpreting individual organisms in a teleological manner. For if an organism is attributed intrinsic value due to its teleological nature, this then implies all things that could be seen as possessing a *telos* must also be granted the same consideration. In this way, certain non-biological objects would also have to be granted this value. As Helgie Kuhse writes:

This understanding [of organisms as being teleological] has the awkward result that we will have to attribute moral standing not

¹⁹Callicott [2] p. 313.

²⁰Leopold [10] pp. 24–25.

²¹Vilkka [21] p. 73.

only to all living things, but to non-living things as well.²²

To emphasise the absurdity of this conclusion, imagine a missile has just been launched with its goal to destroy a particular city. If one were to divert this missile from its intended destination, and thus stopping it achieving its goal, this can be seen, similarly to biological organisms, as interfering with its *telos*.²³

The question thus posed is this: is it acceptable to reasonably think that the action of disrupting a missile from achieving its goal is morally reprehensible? If we are to take the teleological argument seriously, the simple answer is that there ought to be moral consideration attributed in not interrupting the trajectory of this missile, given it has been granted intrinsic value because of its teleological nature. While not the definitive point I wish to make, this above example provides a suitable starting point in drawing incredulity for understanding the teleological argument as a strong foundation for an environmental ethical theory.

The above missile example does not completely undermine the teleological argument because some important theoretical distinctions have since been made by biocentrists in their defence of this argument. J. L. Arbor writes that using the example of the missile to exhibit the weakness of using the teleological argument as a foundational premise of biocentrism and ecocentrism results from a misinformed interpretation. For to state non-biological objects possess a *telos* is to anthropocentrically perceive them to do so; biological organisms stand separate from these because they “have end-states which are not decided by human beings”.²⁴ After all, if we are to remind ourselves of Taylor’s outline of biocentric individualism, the *telos* of biological organisms exists whether or not it is anthropocentrically perceived. Though this clarification goes some way to aid the teleological argument, in actuality this response only creates further problems.

Arbor’s distinction relies on the condition that there exists a complete dividing line between non-biological objects and biological organisms based upon whether the entity under consideration has an end-state anthropocentrically decided by human perception. While this is clear when we take the missile example into account, there still exist non-biological objects that continue to blur this line. That which I have in mind refers to radioactive decay, a non-

²²Kuhse [9] p. 147.

²³Ibid.

²⁴Arbor [1] p. 337.

biological occurrence possessing a *telos* that would transpire whether or not it is anthropocentrically perceived. Radioactive decay involves a compound that “has an unstable combination of nucleons and emits radiation in the process of regaining stability”.²⁵ Accordingly, the good of the radioactive substance here is in maintaining stability; to disrupt this process, in the same manner as a biological organism, would be to harm the good of the object in interrupting its *telos*. Importantly, however, given the above distinction made by Arbor, radioactive decay occurs whether or not humans are there to readily observe it happening. In which case, radioactive decay, according to the above argument made by the biocentrists, ought to be granted intrinsic value.

It is clear that, under biocentric individualism, the teleological argument’s intended use is to extend the scope of intrinsic value to all biological organisms to counter anthropocentric thought. However, given the above result that radioactive decay ought to be granted moral consideration, the teleological argument broadens the range of intrinsic value to non-biological objects that would appear absurd to do so. Of course, one could simply accept from this outcome that radioactive decay has intrinsic value; however, this is an argument that the biocentrist must then stringently defend, an argument that would, in my mind, uncomfortably stretch this philosophical account beyond its principled aim. Perhaps the more sensible route would be to greater distinguish why only the *telos* of biological organisms is to be granted intrinsic value and not that of non-biological objects – a distinction that the teleological argument certainly does not make. Consequently, this argument does not provide a firm basis for the biocentric individualist to argue from, given the argument fails to achieve precisely what is desired from the biocentric outlook.

A closer look at the language used within the teleological argument can provide a form of analysis for why this reasoning leads to absurd conclusions; this moves on to my second point. In understanding biological organisms as being teleological, they are recognised as being benefited or harmed if it interferes with their good; accordingly, moral consideration is to be granted in this way. However, within the framework of applying the teleological argument to all biological organisms, such usage of the words ‘benefited’ and ‘harmed’ become entirely metaphorical when taken in such a broad context. Such utilisation of this vocabulary ought to be therefore treated with frank scepticism, when recognising the emotive consideration expected from the usual connotations of applying these words to describe particular organisms.

²⁵Dawkins [5].

To make this point clearer, let us understand this argument in relation to a vegetable, namely a potato. The teleological argument, under biocentric individualism, would state that to interfere negatively with the good of a potato crop is to cause ‘harm’ towards it. Pulling the potato out of the ground before it is fully grown, for example, can plausibly imply that ‘harm’ has been caused to this vegetable’s *telos* of fully maturing; the question then is, does this vegetable now demand direct moral consideration? I answer in this way: the biocentrist wrongly attributes moral consideration to this interpretation of ‘harm’, because they mistakenly empathise with this non-sentient organism in the same way one would to a sentient being. Only metaphorical ‘harm’ is caused to the potato, because the vegetable is not cognisant of its well-being; this interpretation of ‘harm’ is purely anthropomorphised. As such, moral consideration should not be demanded in the same way as we would towards a sentient being. As Kuhse summarises:

Philosophers are not wrong when they suggest that non-sentient living things can be benefited and harmed, but that they are wrong when they assume that morality has to do with benefits and harms in this broad sense.²⁶

Similarly, Joel Feinberg has strongly argued that by using the word ‘flourishing’ to describe a plant achieving its goal, “nothing is gained by twisting the botanical metaphor back from humans to plants”.²⁷ The use of metaphoric language here is misleading, devoid of meaningfulness regarding moral consideration in the way the biocentric outlook intends.

Given biocentric individualism provides the most explicit use of the teleological argument, the above critique necessarily proceeded with some detail. As will be now shown in relation to biocentric pluralism, the same above points resonate particularly troublingly; however, given these two perspective’s different focus of the attribution of intrinsic value, there exist separate considerations to enquire upon. As previously quoted, Sterba writes that species evolve, split and bud off. While recognising the *telos* of a species in flourishing and surviving may seem straightforwardly identifiable, in fact Sterba’s analysis presents greater complications when attempting to attribute intrinsic value.

If species are to be considered teleological entities, in that they have a good of their own like the individual organisms that comprise them, this assumes clear

²⁶Kuhse [9] p. 147.

²⁷Feinberg [7] pp. 54–55.

distinctions between groups of biological organisms in granting intrinsic value. Yet the taxonomy of doing so is incredibly complex; as Charles Darwin recognised, there exist no fixed essences of a species except the individual biological organism. Darwin saw the term species “as one arbitrarily given for the sake of convenience to a set of individuals closely resembling each other”.²⁸ This presents a problem for Sterba’s argument; for a consequence of this is that it becomes very difficult to categorise distinct species in granting moral consideration if individuals are constantly changing in form. Any identified species that would have naturally died out as a result of continual mutations or adaptations of individual organisms would seemingly demand the moral consideration to be kept in existence due to their apparent teleological nature. Such a scenario ostensibly envisions an environmental policy based upon maintaining a static state of evolution; for if a genetic change threatens the *telos* of one species towards a different classification, this implies biological restrictions would have to be imposed to prevent this species from dying out in maintaining its flourishing and survival. This outcome I argue represents the illogicality of interpreting species as teleological entities. Even if the biocentric pluralist argued that part of a species’ innate *telos* is to eventually become extinct, this then presents an outward contradiction with its advocacy of flourishing and survival laid out in this environmental ethic’s basic premise. *Prima facie* this is an inconsistency that needs to be dealt with, for no being is infallible in the knowledge of gauging the lifespan of an entire species.

Similarly to biocentric individualism, the language used within the biocentric pluralist approach that utilises the teleological argument does not aid in resolving any quandaries with attributing intrinsic value to whole species. To state that a species can be ‘benefited’ or ‘harmed’ is of the same incorrect perspective in referring to all biological organisms in the broad sense that implies moral consideration. Any beneficiary or suffering experienced by a species is again entirely metaphoric; species do not suffer, for instance, because “suffering is an individual matter, not a property of species”.²⁹ Thus, using the teleological argument to account for larger environmental entities still posits pertinent issues that need to be resolved if this argument is to remain the fundamental premise of biocentrism as a whole.

While ecocentrism falls into very similar pitfalls as biocentrism, the consequences for using the teleological argument within this environmental ethic

²⁸Darwin [3] p. 53.

²⁹Dawkins [4] p. 28.

perhaps pose even graver outcomes. For the ecocentric focus on the *telos* of whole ecosystems, regarding their stability and composition, is not only to presuppose ecosystems as being able to reach an equilibrium, but also, due to their perceived teleological nature, there is moral consideration to be granted in maintaining this. Yet, this is an argument flawed from the outset. As Sterba has pointed out, ecosystems do not tend towards stability or harmony; this is “an ecology that is now widely challenged”.³⁰ Ecosystems, like species, are constantly changing, and so to try and maintain stability, as seen from a given perspective of the biotic good, would be ultimately too peremptory and paradoxically the opposite of the holistic mindset that ecocentrism demands.

Furthermore, if only to supplement the above point, theoretical attempts to defend ecocentrism in this manner have been largely discredited; James Callcott, as well as Devall & Sessions, have been criticised for advocating a land ethic that would require the greatest moral consideration to be given to the stability of the biotic community. As Tom Regan has argued, such a scenario would allow for the individual to be sacrificed for the greater biotic good, creating a condition of “environmental fascism”.³¹ With ecosystems existing in constant flux, there are no means for saying what the biotic good can ever be, and thus Regan’s argument represents, as a *reductio ad absurdum*, is a strong reason why ecosystems should not be granted intrinsic value due to their perceived teleological nature.

This is without mentioning once more the mistaken use of language used to describe the fluctuations of ecosystems. For if the biotic good can never be established, there exist no concrete means of stating whether an ecosystem is being ‘benefited’ or ‘harmed’ from internal or external inputs. The only meaningful way an ecosystem could be stated to be negatively or positively affected is through understanding the well-being of the individual organisms that comprise it. As Varner has written, “an ecosystem is valuable only as a means to satisfying individual interests”.³² To state otherwise, in reference to an ecosystem as a whole being, is to again only utilise metaphoric description that is empty of moral deliberation.

³⁰Sterba [15] p. 369.

³¹Regan [12] p. 362.

³²Varner [19] p. 11.

5 Conclusion

Having systematically critiqued the use of the teleological argument throughout biocentrism and ecocentrism, I have argued and hoped to have persuaded the reader that the teleological argument, with the above two reasons I have given, is far from the sturdy premise needed to base these two environmental ethical theories upon. As I have limited this discussion to being a polemic, I shall not here recommend alternative theoretical bases to be used instead of the teleological argument; nor do I advocate that anthropocentrism is the only viable ethical option. However, what I have aimed to achieve is stressing that if biocentrism and ecocentrism are to be considered seriously in alternation to anthropocentrism, a far firmer premise than the teleological argument is desired. Otherwise, it is my opinion that both environmental ethical theories are ultimately subverted, unless they are able to reconcile or respond to any of the fraught issues raised within this paper.

References

- [1] Arbor, J.L. (1986) 'Animal Chauvinism, Plant-Regarding Ethics, and the Torture of Trees', in *Australasian Journal of Philosophy* 64(3): 355–339.
- [2] Callicott, J.B. (1980) 'Animal Liberation: A Triangular Affair', in *Environmental Ethics* 2(4): 311–318.
- [3] Darwin, C. (1859) *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray.
- [4] Dawkins, M.S. (1980) *Animal Suffering: The Science of Animal Welfare*. London: Chapman & Hall.
- [5] Dawkins, M.S. (1998) 'Decay and Half Life', *Integrated Environmental Management, Inc.* [Online] Available at <http://www.iem-inc.com/prhlfr.html>
- [6] Devall, B. & Sessions, G. (1985) *Deep Ecology: Living as if Nature Mattered*. Layton, UT: Gibbs M. Smith Inc..
- [7] Feinberg, J. (1974) 'The Rights of Animals and Unborn Generations', in W. T. Blackstone (ed.), *Philosophy and Environmental Crisis*. Athens, GA: University of Georgia Press.

- [8] Goodpaster, K.E. (1978) 'On being Morally Considerable', in *Journal of Philosophy* 75(6): 308–25.
- [9] Kuhse, H. (1985) 'Interests', in *Journal of Medical Ethics* 11(3): 146–149.
- [10] Leopold, A. (1949) *A Sand County Almanac*. New York, NY: Oxford University Press.
- [11] Norton, B.G. (2003) 'Environmental Ethics and Weak Anthropocentrism', in A. Light & H. Rolston III (eds.), *Environmental Ethics*. Oxford: Blackwell Publishers Ltd.
- [12] Regan, T. (1983) *The Case for Animal Rights*. Berkeley, CA: University of California Press.
- [13] Rowe, S.J. (1994) 'Ecocentrism: The Chord that Harmonises Humans and Earth', in *The Trumpeter* 11(2): 106–107.
- [14] Scruton, R. (2012) *Green Philosophy: How to Think Seriously About Our Planet*. London: Atlantic Books.
- [15] Sterba, J.P. (1995) 'From Biocentric Individualism to Biocentric Pluralism', in *Environmental Ethics* 17(2): 191–207.
- [16] Sterba, J. P. (1998) 'A Biocentrist Strikes Back', in *Environmental Ethics* 20(4): 361–76.
- [17] Taylor, P. (1983) 'In Defense of Biocentrism', in *Environmental Ethics* 5(3): 237–243.
- [18] Taylor, P. (2011) *Respect for Nature: A Theory of Environmental Ethics*, 25th Anniversary Edition. Princeton, NJ: Princeton University Press.
- [19] Varner, G. (1998) *In Nature's Interests: Interests, Animal Rights, and Environmental Ethics*. New York, NY: Oxford University Press.
- [20] Varner, G. (2002) 'Biocentric Individualism', in D. Schmidtz & E. Willott (eds.), *Environmental Ethics: What Really Matters, What Really Works*. New York, NY: Oxford University Press.
- [21] Vilkkka, L. (1997) *The Intrinsic Value of Nature*. Atlanta, GA: Rodopi B. V. Editions.
- [22] Worster, D. (1980) 'The Intrinsic Value of Nature', in *Environmental Review* 4(1): 43–49.

On the Relation Between the Form of Beauty and *Eudaimonia* in Plato's *Symposium**

Carl-Otto Frietsch
University of Warwick

1 Introduction

This essay aims to critically assess the relation between the Form of Beauty and *eudaimonia* (human flourishing). I will argue – firstly – that the purpose of Plato's *Symposium* is to outline how to achieve *eudaimonia* through contemplation of this Form, rather than to explore the best way of engaging in interpersonal relationships with others. I thus agree with Frisbee Sheffield that the criticism leveled at Plato by Gregory Vlastos – which accuses his dialogue of cold-hearted egoism for giving a too instrumental account of *eros* and interpersonal relationships – is essentially misguided. However, I disagree with Sheffield's further claim that this contemplation is a lonesome one that does not involve other people. I therefore intend to argue – secondly – that a life of contemplating the Form (a life of *eudaimonia*) necessarily involves being with others due to the dialogue's characterization of *eros* as a *daîmon* that travels back and forth between the human and divine realms. I conclude on this basis that contemplation of the Form – the relation between the Form of Beauty and *eudaimonia* – is an essentially social activity.

I will begin by explaining how some of the debate has incorrectly understood the dialogue as searching for the best way of loving other people, after which I shall turn to a rival reading of it that instead interprets its aim as being the pursuit of the appropriate method for attaining *eudaimonia*. Finally, I intend to offer an account of why a life of *eudaimonia* is an inevitably social – rather than solitary – life.

*Delivered at the BUPS Autumn Conference 2013 on 23–24 November at the University of Manchester.

2 The misguided debate

Plato's *Symposium* has for a long time been understood as a discussion about love between individuals. The dialogue is set up as a symposium – a kind of drinking party – dedicated to praising 'eros', which may be translated as 'passionate love' or 'desire'. The participants take turns to deliver this praise by way of speeches, which all aim to elucidate eros' nature and function. Central to the text is an argument given by Socrates on behalf of the priestess Diotima. In this so-called 'ascent passage', Socrates describes a hierarchy of beautiful objects. The lover begins by loving the beauty of one body,¹ only to then realize that beauty is similar in all bodies.² The next step is to love the beauty of minds, followed by practices and laws³ and finally the Form of Beauty itself.⁴ This has been interpreted to mean that eros' purpose is to unify one with an abstract and universal object, whilst other entities and people are only desirable insofar as they possess some of the beauty that this Form embodies. Gregory Vlastos criticizes Plato on this basis; accusing him of saying that eros is always directed towards the *beauty* of some object rather than the object itself. He writes that the argument is not one expressing love for individuals, but that "(...) we are to love the persons so far, and only insofar, as they are good and beautiful".⁵

Many scholars have inherited Vlastos' interpretation. One of them is Martha Nussbaum, who argues that the kind of eros one has for a unique individual is lost as a consequence of that the ascent conceives of all beauty as identical.⁶ On her reading of the *Symposium*, Plato accommodates this more 'intuitive' conception of eros through the speech of Alcibiades.⁷ Plato's conclusive view on eros is – according to her – a battle between this kind and that of the ascent, where one is forced to choose one at the expense of the other. I think this reading of the dialogue draws upon a contemporary understanding of love with its interpretation of Alcibiades' speech, which I further believe threatens to situate the discussion of the dialogue within a context that alienates it from

¹Plato [3] 210a.

²Ibid. 210b.

³Ibid. 210c.

⁴Ibid 211c.

⁵Vlastos [6] p. 31.

⁶Nussbaum [2] pp. 117–8.

⁷Nussbaum [1] p. 198.

its own content. I fear that such a direction might obscure whatever message Plato is trying to convey, insofar as he is attempting to convey anything at all.

The following section argues that the *Symposium* does not have the kind of sexual *eros* that one would ordinarily associate with interpersonal relationships as its focus, but is rather seeking to show that having *eros* for *wisdom* is the means by which one is ultimately able to contemplate the Form of Beauty, which consequently allows one to achieve *eudaimonia*. The dialogue is therefore to be viewed primarily as exploring a method to achieve this supreme way of being, rather than seeking the best way of loving others.

3 *Eudaimonia in the Symposium*

It is in the elenchus with Agathon that Socrates brings up *eudaimonia* as an explicit topic of conversation in the dialogue. After Agathon concludes that *eros* is the desire for beautiful things, Socrates asks him whether he thinks that “(...) things that are good are also beautiful”, to which he agrees.⁸ On this basis they both assert that insofar as *eros* makes one desire beautiful things, one also desires good things. Socrates further develops this point in his re-telling of his conversation with Diotima, where he is asked what he who desires beautiful things “(...) get when beautiful things become his own”.⁹ He does not know the answer at first, but Diotima helps him by substituting the word ‘beautiful’ for ‘good’.¹⁰ Socrates then responds immediately that one will be happy when good things become one’s own, which they both agree puts an end to the question of why one desires good things.¹¹ When connected with Socrates’ previous exchange with Agathon, beautiful things are good things that are desired to the extent that they make people happy. Socrates subscribes to this connection by agreeing with Diotima when she proposes, “(...) that the only object of people’s love is the good”.¹²

This part of the conversation culminates in the agreement that “(...) love is the desire to have the good forever”¹³, which I think makes a completely

⁸Plato [3] 201c.

⁹Ibid. 204d.

¹⁰Ibid. 204e.

¹¹Ibid. 205a.

¹²Ibid. 205e–206a.

¹³Ibid. 206a.

different interpretation of the *Symposium* possible. Importantly, *eros* is here extended beyond its sexual aspect so as to be the possible aim of anything that one may concern oneself with in a potential search for happiness.¹⁴ The previous speeches all implicitly situate *eros* as relating to some such specific pursuit. For Phaedrus, *eros* is “(...) the most effective in enabling humans to acquire courage and happiness”.¹⁵ For Pausanias, the lover is making the boyfriend wise and good (184d). In Eryximachus’ case, “(...) the love whose nature is expressed in good actions, marked by self-control and justice (...) is the source of all our happiness”.¹⁶ As for Aristophanes “(...) our human race can only achieve happiness if love reaches its conclusion, and each of us finds his loved one and restores his original nature”.¹⁷ For Agathon, “(...) all good things came to gods and humans through the love of beauty”.¹⁸ This final claim is also serving as the ground for the elenchus with Socrates, which is then linked to Diotima and Socrates’ previously mentioned agreement that *eudaimonia* is the proper object of *eros*.

Insofar as *eros* is ultimately aiming towards *eudaimonia*, what is its relation to beauty? Diotima says that “(...) love’s function is giving birth in beauty both in body and in mind”,¹⁹ which points towards reproduction. This is because reproduction is as close as mortals can get to immortality, which must be desired along with *eudaimonia* if the real end of *eros* is to be the permanent possession of the good.²⁰ However, such reproduction cannot take place in ugliness; it requires beauty in order to happen at all.²¹ What follows is that *eros* – in virtue of having the permanent possession of the good as its genuine end – must also desire beauty in order to possess the good in this way. Through *eros* one thus becomes pregnant in body and in mind, which in beauty leads to the production of different kinds of good things. A man pregnant in body might pursue beautiful women in order to produce what he takes to be eternal

¹⁴Ibid. 205d.

¹⁵Ibid. 180a–b.

¹⁶Ibid. 188d.

¹⁷Ibid. 193c.

¹⁸Ibid. 197b.

¹⁹Ibid. 206b.

²⁰Ibid. 207a.

²¹Ibid. 206d.

happiness in the form of remembrance, by way of creating children.²² He who is more pregnant in mind than in body might instead seek out a beautiful mind in order to give birth to education, which would be his desired end.²³ The specific pursuit of happiness therefore determines one's search for beauty, which is sufficient to shine new light upon the role that the 'ascent passage' has in the dialogue. On this reading it does not just explore how one can give birth to good things in beauty through any kind of *eros*, but *which* kind of *eros* is able to do so to the *highest good there is (eudaimonia)*. Frisbee Sheffield argues that the ascent shows "(...) an *eros* for wisdom in action", which in this context means that one is pursuing an understanding of beauty so extensive that it gives birth to *eudaimonia*.²⁴ This is – of course – the reason why the Form is pursued. I support this reading for the following reason. Before his re-telling of the conversation with Diotima, Socrates criticizes the other speakers for "(...) giving the *appearance* of praising love" (my emphasis), rather than actually doing so.²⁵ This kind of mimetic praising is later referred to by Diotima when she says "(...) that it's only in that kind of life, when someone sees beauty with the part that can see it, that he'll be able to give birth not just to images of virtue, but to true virtue".²⁶ The 'part that can see it' is here what allows one to gaze "(...) at the object with the right part of himself",²⁷ which in the translation by Christopher Gill is footnoted by him as being "(...) the rational mind, when properly prepared to understand the Form".²⁸ He points towards Plato's *Republic*, where one finds that:

(...) A genuine lover of knowledge innately aspires to reality, and doesn't settle on all the various things which are assumed to be real, but with his love remaining keen and steady, until the nature of each thing as it really is in itself has been grasped by the appropriate part of the mind.²⁹

Sheffield maintains that this *eros* for *wisdom* explains the reflection that hap-

²²Ibid. 208e.

²³Ibid. 209b.

²⁴Sheffield [5] p. 128.

²⁵Plato [3] 198e.

²⁶Ibid. 212a.

²⁷Ibid. p. 212a.

²⁸Ibid. p. 80, n. 124.

²⁹Plato [4] 490b.

pens at every level of the ascent, as well as why beauty is portrayed in the instrumental fashion that it is.³⁰ Diotima explicitly says that “(...) if he is to pursue beauty of form, it would be very foolish not to regard all bodies as one and the same” (my emphasis).³¹ The full extent of this understanding is the contemplation of the Form of Beauty itself, which suggests that the ascent is an account of how *eros* for wisdom in the end leads to *eudaimonia*. One has *eros* for wisdom, which requires that one reflect on the common feature of beauty. Reflection upon this beauty culminates in contemplation of its Form, which gives birth to *eudaimonia*.

If the *Symposium* seeks to explicate a method for attaining *eudaimonia* that essentially involves having *eros* for wisdom, then Sheffield’s criticism of Vlastos is warranted.³² It is unproblematic that people are treated as “(...)place-holders of (...) predicates”³³ if *eudaimonia* requires an understanding of beauty that is tantamount to its Form, which through its universality necessitates that one move beyond the beauty found in personal affection. There is no reason to assume that what is meant here is sexual *eros*, because it has already been declared that *eros* can refer to a much wider range of phenomena than that.³⁴ This implies that the debate that stems from reading the *Symposium* as trying to find the correct way of loving others is misunderstanding the dialogue, but it does not mean that relationships with people do not play an important role within it. One of the key aspects of the ascent is the relationship between the guide and the person being guided. What is described as “(...) loving boys in the correct way” amounts to having the lover educate the other in what makes for a good and virtuous way of life.³⁵ On Sheffield’s view, Plato is providing a new perspective on pederastic relationships.³⁶ The kind of exchange-based model that is seen in places throughout the dialogue in the shape of a young boy gratifying a good older man for purposes of “(...)becoming better”³⁷ is replaced by a reflection on beauty that in the process not only provides wisdom, but ultimately human flourishing as well. This would explain why Socrates

³⁰Sheffield [5] p. 128.

³¹Plato [3] 210b.

³²Ibid. p. 129.

³³Vlastos [6] p. 26.

³⁴Plato [3] 205d.

³⁵Ibid. 211b.

³⁶Sheffield [5] p. 131.

³⁷Plato [3] 185a–b.

refuses the invitation of the beautiful Agathon to lie down beside him and – by way of contact – ‘share’ his wisdom with him.³⁸ It would also explain why Socrates chose to ignore Alcibiades’ attempts to gratify him in exchange for knowledge.³⁹

A further question to ask is whether one can have relationships with others in a life of *eudaimonia*. Does a life of contemplating the Form of Beauty involve other people at all? Sheffield does not think so.⁴⁰ She thinks that if Plato required the person at the top of the ascent to do something such as generate virtue in other souls, then he has been “(...) spectacularly unclear in his exposition”. I disagree with this claim. The next section outlines an argument for why contemplation of the Form fundamentally involves other people.

4 *Eudaimonia* and relationships

This far I have argued – like Sheffield – that contemplation of the Form is synonymous to giving birth to *eudaimonia* through eros for wisdom.⁴¹ My further claim is that a life of contemplating the Form (a life of *eudaimonia*) itself features eros as a vital component. Sheffield does not say that this kind of life stands in such a relation with eros, but I think that her analysis further entails it. She holds that eros for wisdom manifests itself as reflection, but she does not explicate the relation between this reflective state and the contemplative element that is part of the contemplation of the Form. Apart from *what* is being reflected or otherwise contemplated upon, I claim that there is no difference between the two cases. In both instances there is the moment of giving birth in beauty, which Diotima maintains to be the proper function of eros.⁴² If one finds Diotima’s story to be credible therefore, it is only reasonable to agree that contemplation of the Form also involves eros. Given the extent to which Sheffield’s own account relies on the accuracy and significance of the words of Diotima, I do not think that he has any reason to object to this inference.

If one accepts that contemplation of the Form involves eros in this way, then

³⁸Ibid. 175d–e.

³⁹Ibid. 217a–c.

⁴⁰Sheffield [5] p. 133.

⁴¹Ibid. p. 128.

⁴²Plato [3] 206b.

I hold that there is a case for saying that a life of contemplating the Form is not at all the solitary practice that Sheffield says it is. In order to understand why this is so, our attention must once again turn to a specific part of the elenchus between Diotima and Socrates. Here, the dialogue's portrayal of *eros* is changed from being a God to that of a *daïmon*.⁴³ This 'great spirit' is between the mortal and the immortal; delivering messages from humans to Gods and back again. It is claimed that this *daïmon* is the child of Resource and Poverty,⁴⁴ for which reason the resources he obtains are always drained away.⁴⁵ He is neither with nor without them, but is always in a state of need;⁴⁶ he is between of wisdom and ignorance.⁴⁷ What follows is that contemplation of the Form – insofar as one has *eros* in this state – is also in between of wisdom and ignorance since *eros* is a *daïmon*. It must be a contemplation that is in a perpetual need for wisdom. On this basis it is further possible to infer that a life of *eudaimonia* is one of needing and desiring to have wisdom. If this is so, then perhaps Plato offers us a clue to what this life is like with the character he depicts as loving wisdom the most in the text. It is consequently difficult to deny Socrates this place, and especially so in the *Symposium*. In fact, Plato here greatly emphasizes Socrates' unique *eros* for wisdom and his desire for knowledge. Towards the end of the dialogue it is described how he stays up after the party is over and continues to debate throughout the night, only to stop once the last person awake has finally fallen asleep.⁴⁸ However, we are also told how he then immediately left the party and went on about his day "(...) as he did at other times, and only then in the evening went home to bed", which strongly suggests that he tenaciously kept on pursuing wisdom far more extensively than even the most highly esteemed thinkers of his time. To this I add that Socrates – even by Sheffield's lights – is "(...) a seeker after wisdom *par excellence*"⁴⁹. If one therefore asserts that Socrates has attained a life of *eudaimonia* – where such a life essentially involves a constant desire for wisdom, as I argue – then the way in which he pursues wisdom is going to be how he contemplates the Form. Does Socrates then pursue wisdom in a soli-

⁴³Ibid. 202d–e.

⁴⁴Ibid. 203b.

⁴⁵Ibid. 203e.

⁴⁶Ibid. 203d.

⁴⁷Ibid. 203e.

⁴⁸Ibid. 223b–e.

⁴⁹Sheffield [5] p. 136.

tary way, as Sheffield suggests? Or does he – perhaps more than anyone ever has throughout the history of philosophy – pursue a *dialogue with others* due to his desire to know and be wise? I think the answer to this question is clear. A flourishing life of contemplating the Form *just* is the Socratic dialectic. This explains why Diotima – the character in the dialogue most likely to have attained *eudaimonia* – ‘teaches’ Socrates despite ‘possessing’ wisdom,⁵⁰ which is also a problem that Sheffield chooses not to explore further in her analysis.⁵¹ If the way to live a life of *eudaimonia* is the same as pursuing a dialogue in search of wisdom, then I think that this offers an important clue to the reason why Plato writes in dialogue form. Socrates becomes ‘immortal’ in a search for wisdom that goes on forever throughout Plato’s books.⁵² If the permanent possession of the good is this perpetual pursuit for knowledge, then he has definitely secured it for Socrates. Perhaps the reason why the *Symposium* is not written in the standard format is because Plato needed a way to emphasize how important the dialectic is if one wishes to praise *eros* properly. After all, Socrates’ speech is nothing other than a dialogue; it is a speech that is not really a speech. This is also true for Apollodorus, who is the actual narrator of the story.⁵³ At any rate, I think that a good case has been provided for the view that a life of *eudaimonia* is not the solitary engagement that Sheffield appears to think that it is. This is because the relation between the Form of Beauty and *eudaimonia* is not a lonesome contemplation that is done on one’s own, but a *dialectical* contemplation that can only be established in the company of other people.

References

- [1] Nussbaum, M. (2001) ‘The Speech of Alcibiades: A Reading of Plato’s Symposium’, in *The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy*. Cambridge: Cambridge University Press.
- [2] Nussbaum, M. (1992) *Love’s Knowledge: Essays on Philosophy and Literature*. Oxford: Oxford University Press.

⁵⁰Plato [3] 206b.

⁵¹Sheffield [5] pp. 134–35.

⁵²Plato [3] 212a.

⁵³Ibid. 174a.

- [3] Plato (1999) *The Symposium*, C. Gill (trans.). London: Penguin Classics.
- [4] Plato (2008) *Republic*, R. Waterfield (trans.). Oxford: Oxford University Press.
- [5] Sheffield, F. (2012) ‘The Symposium and Platonic Ethics: Plato, Vlastos, and a Misguided Debate’ in *Phronesis*, 57(2): 117–41.
- [6] Vlastos, G. (1973) ‘The Individual as Object of Love in Plato’, in *Platonic Studies*. Princeton: Princeton University Press.

Hegel's *Philosophy of Right*: Civil Society and International War

Joel Mason

University of Exeter

Introduction

In Hegel's *Philosophy of Right*, he gives accounts both of relations within civil society, and of relations between states. However, his vision for the two is markedly different – in civil society, Hegel starts with the individual free will, but believes that interdependence between individuals will strengthen civil society and bind it together; in the sphere of international relations, Hegel starts with the state as sovereign, but considers war the only likely outcome given the lack of a higher authority between states. In this essay, I shall argue that his account of the former provides various reasons to believe that his predictions on the latter are unjustified. I shall argue, firstly, that his account can be used in relation to the transnationalisation of civil society, such that his argument not only explains the societal bonds that exist within the borders of one state, but rather suggests that relations between people of different states should be of the same nature, and thus that war between states is not likely. Secondly, I shall argue that his account of the relations between people in civil society provides a logic which can be applied, by analogy, to the sphere of international relations, and which thereby contradicts his claim that war is likely. I therefore hope to show, via these two arguments, that Hegel is unjustified in claiming that war is the likely outcome of international relations.

1 Hegel's Account of Civil Society

Before I put forward these arguments, however, it shall be necessary to very briefly outline Hegel's account of the relations between people in civil society, and particularly the notion of interdependence, on which much of the proceeding arguments rest. I shall also briefly set out Hegel's reasons for arguing that war is the likely result of international relations.

Hegel's notion of civil society is based upon his notion of right, which is in turn founded upon the concept of the free will. Hegel's account throughout the *Philosophy of Right* is of the realisation of freedom, and the free will is the

starting point of his philosophical system; the basic building block upon which the rest is constructed. The free will first actualises itself by externalising itself in objects in the world, and thereby becoming ‘right’.¹ It is thus by appropriating things through the embodiment of externalised will that individuals come to have property.² However, these individual wills, embodied in objects in the world, inevitably come into conflict. This necessitates contracts – and in the mutual recognition involved in the agreement of these contracts individuals truly recognise the other party as a rights-holder and as an individual.³ It is therefore through creating an ever more complex and interwoven web of mutual interdependence that the bonds of society are created and strengthened. This is the realm of ‘abstract right’. Hegel’s account thus far, however, explains only the immediate relations of people to one another; human freedom is, though, only fully realised as part of an “actual ethical order”⁴ – that is to say, within and through the institutions of civil society and the state, such as the free market, the justice system etc. These are the realisation of right and the free will.

Ultimately, however, Hegel does not see a peaceful future for these states. Given that states are “sovereign and independent in relation to others”, Hegel argues, it is impossible to “get beyond an ‘ought’” when considering the obligations imposed upon them by treaties, because they remain forever above the demands of the treaties, and as such the relations between states must remain in a “state of nature”.⁵ The effect of this is that “if states disagree and their particular wills cannot be harmonised, the matter can only be settled by *war*”.⁶ War remains a structural possibility within Hegel’s system. Incidentally, Hegel also considers a number of benefits to arise from the waging of war, however these are not the main reason why he deems war likely, and as such I have not sufficient space to expand upon these in this essay.⁷

¹Hegel [2] §41, p. 57 and §40, p. 56.

²Ibid. §44, p. 60.

³Ibid. §71, pp. 83–84.

⁴Ibid. §153, p. 160.

⁵Ibid. §§330–331, p. 311 and §333, p. 312.

⁶Ibid. §334, p. 313.

⁷Various advantages are discussed in § 324, principally focusing on the idea that war serves as an important reminder of the contingency and finitude of life, property and relations. Ultimately, Hegel suggests that “stagnation would be the product of prolonged, let alone perpetual, peace” (Hegel [2] §324, pp. 305–308). However, Hegel writes that “actual wars require some other justification”, and points readers to the points already made (ibid. §324, p. 307).

2 The Transnationalisation of Civil Society

My first argument contra Hegel's claim that international relations are likely to end in war is based on the idea of what Jaeger calls the "transnationalisation of civil society".⁸ Civil society, as conceived by Hegel, brings together (although not as fully as in the state, which is the final moment of ethical life, though in fact is presupposed by civil society) the particular and the universal in its three domains: the system of needs (i.e. the market), the administration of justice (i.e. the legal system), and the police and corporations, by which is meant public policies.⁹ I shall argue that all of these elements of civil society have transnational equivalents, particularly (and increasingly) in the modern day.

Civil society is characterised by the system of needs, i.e. the market, through which the bonds of mutual interdependence are strengthened, as explained above. One of the reasons, perhaps, that Hegel confined this process to within the borders of individual states was his view that "independent states are principally wholes which achieve satisfaction internally" – that is to say, that states are largely self-sufficient entities.¹⁰ However, even if this was the case when Hegel was writing, the dramatic increase in the volumes of world trade which has occurred in the almost two centuries following the publication of the *Philosophy of Right* surely suggests that this is no longer the case. Nonetheless, Hegel appears to give indications at certain points that transnational interdependence through the market may have existed: he admits that "a state... , through its subjects, has many-sided relations" which may be damaged by war, and at another point argues that a state's industries are "dependent on conditions abroad".¹¹ Most explicitly, he suggests that civil society, through the mechanism of the market, strives to create "commercial connections between distant countries and so relations involving contractual rights".¹² Jaeger's argument that such activity "extend[s] civil society itself beyond the nation-state" appears convincing, and importantly, it happens "as a result of the states' own making".¹³

⁸Jaeger [3] p. 509.

⁹Hegel [2] §188, p. 186.

¹⁰Ibid. §332, p. 312.

¹¹Hegel [2] §334, p. 313 and §236, p. 217.

¹²Ibid. §247, p. 222.

¹³Jaeger [3] p. 515.

Equally, regarding the second domain of civil society – the legal/justice system – it appears that there now exist transnational institutions which fulfil the roles envisaged by Hegel. For instance, the UN Declaration of Human Rights sets out universal rights, and bodies such as the International Court of Justice and International Criminal Court have been established as transnational executors of justice.

Finally, as Jaeger points out, it follows that if the market extends beyond the boundaries of the state, then “the negative externalities of the market . . . are also transnationalised”,¹⁴ therefore numerous international bodies have been established to deal with these problems on a transnational basis. For instance many of the UN agencies such as the UN Development Program and the World Food Program, as well as other aid agencies, work to counter education, food and health problems globally. Similarly, international agreements on issues such as environmental protection and arms trade attempt to deal with the problems created by the market in relation to these issues. These correspond to the final domain of civil society – the police and corporations, which aim to mitigate the worst effects of the market system and provide a basic level of welfare.¹⁵

In his criticism of Jaeger’s argument, Brooks dismisses Jaeger’s claims, arguing that for his argument to hold, “we would have to suppose the main characteristics of civil society . . . are present in the international sphere in some form . . . this is clearly not the case”.¹⁶ What I have argued is that these main characteristics – the interdependence of market relations, a legal/justice system, and a system of public policy – *are* in fact present in the international sphere.

What I hope to have shown using this first argument is that there is no reason why civil society, and the bonds of mutual interdependence which it creates through the market system, should be confined to within the borders of a state. Instead, it seems conceivable that civil society could transcend state boundaries, and thereby bind together not just the people of one nation, but people across the world.

¹⁴Ibid.

¹⁵Hegel [2] §§244–245, pp. 221–222.

¹⁶Brooks [1] p. 512.

3 The Logic of Civil Society and its Application to International Relations

My second argument, then, aims to show that the logic which governs relations between individuals, and most particularly how the notion of abstract right binds individuals together in civil society through bonds of mutual interdependence via the system of needs, can also be applied to relations between states in the international sphere, and hence why war is not the likely outcome of international relations.

As stated above, individuals are bound together in society via the market, through contract and the mutual recognition which this entails. Peperzak argues that this logic is transferable, and thus that “the transition from one individual state to a multiplicity of states repeats on a higher level the transition from one human person to a plurality of persons”.¹⁷ Through contract (treaties) states recognise the sovereignty and individuality of other states. Peperzak writes that “just as the situation of individual persons facing one another necessarily developed into the structure of a contract . . . so does the situation of states ruled by the abstract norm of mutual respect necessarily result in their concluding treaties”.¹⁸ Indeed, at certain points Hegel himself appears to suggest such a comparison, writing that “a state is as little an actual individual without relation to other states as an individual is actually a person without a relationship with other persons”.¹⁹

However, Hegel would argue, of course, that there is a crucial difference between the two situations, which is that states are always sovereign and between them there may be no higher judge.²⁰ In response to this, though, it may be argued that the situation is not so different to that facing individuals within the scope of abstract right. At first, crimes were only overcome and justice restored by revenge, however this is imperfect due to the issue of impartiality. This is the situation that states find themselves in. In order to resolve this contradiction in the individual sphere, Hegel moves on to the notion of morality, ethical life, civil society and the state, and ultimately creates a system in which all are bound together by bonds of interdependence, with institutions overseeing and enforcing this. Yet with regard to the very similar situation faced by

¹⁷Peperzak [5] p. 248.

¹⁸Ibid. p. 255.

¹⁹Hegel [2] §331, p. 311.

²⁰Ibid. §333, p. 313.

states, Hegel proposed no such solution. There is no reason why the importance of state sovereignty should pose any greater a difficulty in resolving this problem than the issue of individual free will did for the case of individual relations – i.e. it may be overcome. Peperzak proposes this line of argument, concluding that “it is conceptually necessary to postulate the actuality of a higher instance, a supranational court of justice insuring fairness towards all people and armed by a legal and executive force that is acknowledged by all nations”.²¹ It follows from Hegel’s logic, he claims, that “sovereign states would sacrifice their absolute independence in order to constitute a worldwide organisation, in which humankind as a whole would be able to recognise the realisation of its practical aspirations”.²² He argues that this process would simply be a “repetition on a higher level of Hegel’s own developments of abstract right and morality”.²³ Vincent, too, argues that given Hegel’s underlying dialectic, and his negative account of international relations, it follows that “a positive state of affairs transcending the clash of nation-states” must emerge.²⁴ He argues that “there is nothing which ties ethics and freedom to the nation state, in fact there is no reason why an international ‘ethical life’ or code of customary conduct should not be formulated according to Hegelian principles”, and thus concludes that “an international order of some kind is logically possible”.²⁵

Again, critics of this idea will point out that it is precisely this proposal suggested by Kant that Hegel was at such pains to reject. Kant argued in favour of the creation of a “federation of nations” which he conceived as being “a contract resembling the civil one and guaranteeing the rights of each [nation]”, which he argued should include all nations and would “lead to perpetual peace”.²⁶ However, there are again grounds for arguing that Hegel’s own logic appears to push him towards accepting this idea. Vincent and Plamenatz both argue that the idea of a world-state, or something similar, fits in with Hegel’s idea of Universal Spirit or Mind, as part of his account of the general progress of history. Universal Spirit or Mind elevates itself and achieves ever greater consciousness throughout history, but then Hegel seems to suggest it will stop short of what appears to be its conclusion, of coming together

²¹Peperzak [5] p. 258.

²²Ibid. p. 260.

²³Ibid. p. 261.

²⁴Vincent [7] p. 203.

²⁵Ibid. p. 196 and p. 203.

²⁶Kant [4] pp. 115–117.

in a complete whole. Vincent writes that “the state, with regard to the development of Spirit, is still but a stage characterised by imperfection, finiteness, incompleteness and contingency”.²⁷ However, given that Objective Spirit or Mind (the moment in which discussion of political philosophy occurs) can give way to Absolute Spirit or Mind, this suggests that “the state, by definition, is not the highest realisation of Spirit”.²⁸ Similarly, Plamenatz suggests that “it seems to follow from Hegel’s conception of Mind or Spirit that the fully rational and free community must be a world-state”²⁹ since this would represent the full realisation of the universal will.

4 Conclusion

Ultimately, then, I hope to have shown that Hegel’s prediction that war is the likely outcome of international relations cannot be justified. I hope to have demonstrated how Hegel’s own logic contradicts this claim – firstly, due to the possibility of the transnationalisation of civil society, such that the bonds between people are strengthened not only within the state, but also across borders; and secondly, because of the analogy which may be drawn concerning relations between people and relations between states. The conclusion which I have attempted throughout to establish via these arguments is that Hegel’s arguments regarding the likelihood of war cannot be justified. Instead, his arguments which suggest mutual interdependence will strengthen the bonds of civil society should also provide grounds for believing the same will obtain with regard to international relations.

References

- [1] Brooks, T. (2004) ‘Hegel’s Theory of International Politics: A Reply to Jaeger’, in *Review of International Studies* 30(1): 149–52.
- [2] Hegel, G.W.F. (2008) *Outlines of the Philosophy of Right*, T.M. Knox (trans.). Oxford: Oxford University Press.

²⁷Vincent [7] p. 202.

²⁸Ibid.

²⁹Plamenatz [6] p. 148.

- [3] Jaeger, H. (2002) ‘Hegel’s Reluctant Realism and the Transnationalisation of Civil Society’, in *Review of International Studies* 28(3): 497–517.
- [4] Kant, I. (1988) ‘To Perpetual Peace: A Philosophical Sketch’, T. Humphrey (trans.), in T. Humphrey (ed.), *Perpetual Peace and Other Essays*. Indianapolis, IN: Hackett Pub. Co.
- [5] Peperzak, A. T. (1994) ‘Hegel *contra* Hegel in his *Philosophy of Right: The Contradictions of International Politics*’, in *Journal of the History of Philosophy*, 32(2): 241–63.
- [6] Plamenatz, J (1993) *Man and Society: Hegel, Marx and Engels, and the Idea of Progress, Volume III*. London: Longman.
- [7] Vincent, A. (1983) ‘The Hegelian State and International Politics’, in *Review of International Studies*, 9(3): 191–205.

Are Scientific Revolutions Irrational?*

Sophie Osiecki

University College London

Kuhn's *Structure of Scientific Revolutions (SSR)* presented his thesis on scientific progress.¹ Many philosophers have since criticised his portrayal of science as 'irrational'. The most famous of these was Imre Lakatos who claimed that "scientific revolutions are irrational; a matter for mob psychology" in his paper, 'Falsification and the Methodology of Scientific Research Programmes'.² In this essay I will be focusing on Kuhn's notion of a scientific revolution as presented in *SSR* as this was the work that Lakatos was criticising. I will argue that although Lakatos never fully elucidates what he means by rationality, it is clear that any reasonable interpretation of Lakatos' analysis does in fact render Kuhn's 'revolutions' irrational, but that this is by no means a detrimental characteristic of them.

As the title suggests, Kuhn's *SSR* attempts to outline scientific development. He was responding to the tradition of logical positivists who maintained the view that scientific development was a cumulative process; progressing further to an objective understanding of the world.³ Kuhn rejected this approach and claimed that science develops "without benefit of a permanent fixed scientific truth".⁴

For Kuhn, 'scientific revolutions' are synonymous with his notion of 'paradigm-shifts'. Kuhn gives many historical examples of paradigm-shifts "associated with the names of Copernicus, Newton, Lavoisier, and Einstein."⁵ Broadly, the term 'paradigm' refers to the fundamental set of theories and metaphysical beliefs that an individual or group of scientists hold as characterised by particular 'exemplars' in the field. An example of this is the conception of light within the Newtonian paradigm: Newton's *Opticks* provided grounds for the

*Delivered at the BUPS Autumn Conference 2013 on 23–24 November at the University of Manchester.

¹Kuhn [4].

²Lakatos [6] p. 178.

³Oberheim [7].

⁴Kuhn [4] p. 173.

⁵Ibid. p. 6.

idea “that light was material corpuscles” which in turn implied a ‘mechanico-corpuscular metaphysic’.⁶ The function of a paradigm is to provide the conceptual and methodological underpinnings of scientific endeavour; it articulates what sorts of questions scientists ask, and how they can be answered.⁷

Once a paradigm has been established within a given field, a period of ‘normal science’ follows within which the activity Kuhn calls ‘puzzle-solving’ takes place. During this period, scientists rarely question the paradigm and work on refining their theories. However, anomalies, which are results that deviate from what is expected in accordance with the paradigm, may begin to occur with sufficient frequency as to result in a period of ‘crisis’. There are three possible outcomes of this: (i) a new paradigm will supersede the previous one, (ii) scientists will ignore the anomaly or (iii) the anomaly will become assimilated with the existing paradigm.

Schematically, this process seems quite straightforward. However, Kuhn argues that paradigms are incommensurable with one another; there are no objective standards for their comparison.⁸ This is because paradigms provide the ‘conceptual framework’ which shapes the scientist’s world-view which cannot be compared to ‘the facts’ or a world independent of the scientists.⁹ As a result, Kuhn states that one paradigm may supersede another due to extra-scientific factors; a paradigm shared within a scientific field may become obsolete due to “their members’ conversion to the new paradigm”.¹⁰ Kuhn quotes the physicist Max Planck: “a new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it”.¹¹

Lakatos argues that “Kuhn’s view of scientific revolutions is irrational, a matter for mob psychology”.¹² I will presently examine how he arrived at this conclusion and attempt to ascertain what he meant by ‘irrational’. He states that there are “no rational causes for the appearance of a Kuhnian crisis”; “crisis is a psychological concept, it is contagious panic”. It would seem at this stage

⁶Kuhn [4] p. 10-22.

⁷Bird [1].

⁸Kuhn [4].

⁹Okasha [8] pp. 85-90.

¹⁰Kuhn [4] p. 19.

¹¹Ibid. p. 151.

¹²Lakatos [6] p. 178.

that Lakatos is accusing Kuhn of irrationality on the basis that his account of scientific revolutions takes into account the psychology of the scientists.

Immediately after condemning the Kuhnian notion of ‘crisis’ he goes on to criticise the next stage in Kuhn’s account. He begins by stating: “The new ‘paradigm’ emerges, incommensurable with its predecessor”, a criticism that is essentially congruent with Kuhn’s views, though Lakatos does not develop the criticism in any detail. Lakatos then proceeds to make a number of claims:

- (a) “There are no rational standards for their comparison.”¹³
- (b) “Each paradigm contains its own standards.”
- (c) “The new paradigm brings a totally new rationality.”
- (d) “There are no super-paradigmatic standards for their comparison.”¹⁴

In the first instance, in claim (a), it is unclear exactly what Lakatos means by ‘rational.’ Subsequently, (b) is merely a statement of Kuhn’s exposition of a paradigm. Once more, in (c), Lakatos reintroduces the notion of ‘rationality’ and it remains unclear what his intended meaning is. It is only when we acknowledge (d) that his meaning begins to crystallise. It is notable that (d) closely resembles claim (a), with the exception that ‘rational’ has been replaced with ‘super-paradigmatic.’ Is this what he means by rationality?

It could be argued that in actual fact Lakatos was making two separate claims (a) and (d). If this were the case then firstly he would be claiming that scientific revolutions are irrational because there are no rational standard for the comparison between two competing paradigms. Secondly, he would be claiming that scientific revolutions are irrational because there are no super-paradigmatic standards for the comparison between two competing paradigms. In the first claim, his reasoning is inherently circular; of course scientific revolutions are irrational given that there are no rational standards for their comparison. Furthermore, we must not lose sight of his concluding statement: “Thus, in Kuhn’s view scientific revolutions are irrational. . .” The use of ‘thus’ implies that his previous claims entail this final statement. Overall, it seems reasonable to believe that Lakatos is in fact equating the term ‘rational’ with

¹³By ‘their comparison’ in (a) and (d) Lakatos is referring to two competing paradigms in a time of crisis.

¹⁴Separating Lakatos’ claims in this way is not to suggest that I am taking them as an exposition of a formal argument. On the contrary, I suggest that Lakatos’ argument is highly rhetorical. It is for this reason that I have dealt with it in this way, to try to ascertain his meaning.

‘super-paradigmatic’.

I think at this point it is important to distinguish between the programmes of Kuhn and Lakatos. The aim of Lakatos’ paper was to ‘rationally reconstruct’ scientific development, whereas Kuhn was concerned with a more descriptive and contextualised account of it. Lakatos was aware of this distinction and explicitly condemned Kuhn’s approach: “Kuhn certainly showed that the psychology of science can reveal important and, indeed, sad truths”.¹⁵ This implies that the justification for irrationalism is based on Lakatos’ own pre-conceptions of how scientific development ought to be regarded in the philosophy of science. So what did he mean by rationality? He continues: “But the psychology of science is not autonomous; for the rationally reconstructed growth of science takes place in Popper’s ‘third world’ ... in the world of articulated knowledge which is independent of knowing subjects”.¹⁶ In a footnote Lakatos states that this third world is the “world of ideas” as distinct from the “material world” and “world of consciousness”.¹⁷ The discussion in this paragraph has led us to a narrow interpretation of what Lakatos meant by ‘rationality’; we are drawn to the conclusion that his ‘rationality’ in science is limited to the ‘third world’, which would, broadly, be ‘super-paradigmatic’.

Although this may be the case, Feyerabend argues that: “According to Lakatos the apparently unreasonable features of science occur only in the world of (psychological) thought ... It must be pointed out, however, that the scientist is unfortunately dealing with the world of matter and of (psychological) thought ...”.¹⁸ If one is to accept Feyerabend’s position and Lakatos’ definition of rationality, then it appears that scientific revolutions, as Kuhn describes them, are inherently irrational. Furthermore, Lakatos’ obsession with ‘rationality’ with respect to science is unhelpful in giving us a realistic account of how science actually develops.

In a footnote Lakatos states: “I have not dealt with Kuhn’s and Feyerabend’s claim that theories cannot be eliminated on any objective grounds because of the ‘incommensurability’ of rival theories”.¹⁹ This is a reasonable decision as the introduction of the term ‘objective’ suggests a philosophical position

¹⁵Lakatos [6] p. 179.

¹⁶Ibid. p. 180.

¹⁷Ibid.

¹⁸Feyerabend [2] p. 218.

¹⁹Lakatos [6] p. 180.

known as scientific realism within which it is postulated that our scientific theories correspond with ‘a mind-objective reality’, which leads to difficulties.²⁰ However, there is a striking similarity between the status of such a ‘mind-objective reality’ and Lakatos’ insistence of the ‘third-world’; namely that it is impossible to verify its existence given the constraints that Feyerabend pointed out as discussed above.

In conclusion, Lakatos’ use of ‘rationality’ is self-confessedly found in the ‘third world’, an area Kuhn was not concerned with. It seems reasonable then to account for Lakatos’ views on the basis of his preconceived idea about how scientific development ought to be conceived. He deplored Kuhn’s analysis of the psychology of discovery (i.e. ‘sad truths’²¹) which Kuhn thought was imperative. In this way, Lakatos was right in claiming that Kuhn’s ‘scientific revolutions’ are irrational: however, in comparison with Lakatos’ alternative, they provide us with a way of conceptualising the development of science in its proper context.

References

- [1] Bird, A. (2010) ‘Thomas Kuhn’, in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* [Online] Available at <http://plato.stanford.edu/archives/fall2010/entries/thomas-kuhn>
- [2] Feyerabend, P.K. (1970) ‘Consolations for the Specialist’, in I. Lakatos & A. Musgrave (eds.), *Criticism & the Growth of Knowledge*. Cambridge: Cambridge University Press.
- [3] Hoyningen-Huene, P. (1989) *Reconstructing Scientific Revolutions: Thomas S. Kuhn’s Philosophy of Science*. Chicago and London: University of Chicago Press.
- [4] Kuhn, T.S. (1970) *The Structure of Scientific Revolutions*. 3rd ed. Chicago and London: University of Chicago Press.
- [5] Kuhn, T.S. (1970a) ‘Reflections on my Critics’, in I. Lakatos & A. Musgrave (eds.), *Criticism & the Growth of Knowledge*. Cambridge: Cambridge University Press.

²⁰Sankey [9] p. 74.

²¹Lakatos [6] p. 179.

- [6] Lakatos, I. (1970) ‘Falsification and the Methodology of Scientific Research Programmes’, in: I. Lakatos & A. Musgrave (eds.), *Criticism & the Growth of Knowledge*. Cambridge: Cambridge University Press.
- [7] Oberheim, E. & Hoyningen-Huene, P. (2010) ‘The Incommensurability of Scientific Theories’, in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* [Online] Available at <http://plato.stanford.edu/archives/fall2010/entries/incommensurability>
- [8] Okasha, S. (2002) *Philosophy of Science: A Very Short Introduction*. Oxford: Oxford University Press.
- [9] Sankey, H. (2008) *Scientific Realism and the Rationality of Science*. Great Britain: MPG Books Ltd.

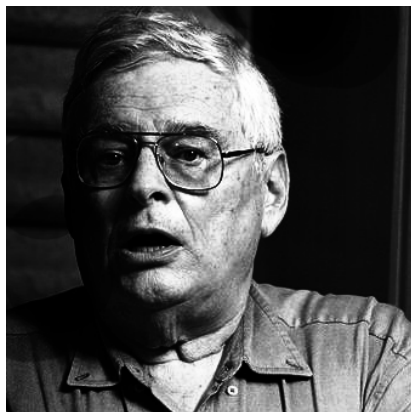
Interview: Jerry Fodor

Rutgers University, New Jersey

What is philosophy and why should one ‘philosophise’?

I think of philosophy as a kind of geography and as a service industry. Often one wants to know what else one will be committed to if one endorses a certain proposition. Answers to such questions can be distinctly unobvious and extremely hard to work out, to say nothing of distinctly embarrassing (as when you find that if you claim that P you must also claim not P , thereby contradicting oneself. There are well known reasons for avoiding self-contradiction if one can). There is no general agreement as to how this process of uncovering connections between propositions should proceed, though many useful models of successful attempts are to be found in mathematics and logic; hence the centrality of both in many philosophical discussions. But there isn't, as far as anyone knows, a general procedure for achieving or figuring out what follows from what. One uses one's experience, one's intuitions, and one's ingenuity, and one thrashes about as best one can. Interesting things sometimes turn up.

As to why you should bother doing philosophy, there isn't any reason that I can think of unless you find it interesting. In my experience, it keeps one up at night and the financial incentives are minimal. But ‘what's-connected-to-what’ questions do present themselves in all sorts of investigations other than the ones with which philosophers have traditionally been concerned, so it's just as well that someone should be trained in pursuing them. One way of getting such training is to work through some of the cases



that have been examined in the history of philosophy. You won't find out what the good is, or what knowledge is, or what an object is, or whether abstract objects are real. But you will find out a lot about the consequences of endorsing one or another proposal, and why so many of the proposals that have been offered from time to time don't work. (In general, philosophy is better at finding out what doesn't work than what does). There is, by the

way, nothing against turning philosophy on itself; ‘what is philosophy?’ is a paradigm of a philosophical question. A flock of answers have been proposed over the years, to all of which philosophers have found persuasive counterexamples.

What is your first memory of philosophy? Whose works did you enjoy?

I don’t at all remember when I first encountered philosophy. Probably it was when I first encountered Lego, which is also about taking things apart and then trying to put them back together. I suppose I’ve always been more comfortable thinking about the kinds of questions that one might reasonably hope to answer by doing thought-experiments and thinking up examples and counterexamples than about the kinds of questions that one can answer by doing experiments in laboratories. Socrates is the prototype. He was very good at, for example, showing people that what they say that they think virtue is, is in conflict with their own intuitions about which acts are virtuous. (Unsurprisingly, he was considerably less good at figuring out what, in fact, virtue is; this is a pattern that iterates throughout the history of philosophy).

Is there any advice you can give to undergraduates aspiring to be philosophers?

I think the best advice about learning how to do philosophy is to learn a lot of things that aren’t philosophy. If you want to do philosophy, you have to know things to do it to. In that sense, philosophy is inherently a second-order discipline. I got into the philosophy of mind in large part by being exposed to empirical work in psychology; and I got into the philosophy of language in large part by being exposed to empirical work in linguistics. In both cases, it seemed that the questions the practitioners worked on were often more interesting than the questions that they said that they worked on when you asked them; their accounts of what they were doing were strikingly bad fits to their practices in doing it. Why, for example, do so many psychologists insist that they are interested in the prediction of a subject’s behavior when it’s perfectly clear that what they really care about is how the subject’s mind works? Why, for that matter, are there so many philosophers who claim to be interested in “analyzing concepts” but are then quite content to leave it to somebody (psychologists maybe) to figure out what concepts are?

Another way to put this: I don’t think there is any such thing as ‘pure’ philosophy, a discipline in which one starts with self-evident (“intuitively evident”), truths and arrives at true conclusions by employing only kinds of inference that are demonstrably sound. From Descartes to Wittgenstein (inclusive) that sort of view of the method of philosophy has been closely tied to the idea that the philosopher’s goal is to refute skepticism: i.e. to refute claims like ‘there

are no such things as tables, chairs, or earthworms'; or claims that we have no reason to believe that what appear to be other people aren't really just mindless machines, and so forth for a number of other theses which, quite rightly, strike the laity as either frivolous or plainly false (or both). I think, rather, that the (dis)confirmation of philosophical theories, much like the (dis)confirmation of empirical theories, is inevitably holistic, which is to say that it's inevitably sensitive to constraints like simplicity, face plausibility, coherence with other theories that are themselves *prima facie* plausible, etc. In particular, I don't think that there is, or even could be, a purely 'analytical' philosophy, the truths of which are all, in some sense, conceptual or semantically necessary. The relevant methodological maxim is not 'only say things that are conceptually or semantically necessary', but 'don't say things that are patently false'. Samuel Johnson was right: If someone tells you that rocks are just 'bundles of ideas', the right thing to do is suggest that he try kicking one. "Thus I refute Bishop Berkeley", Johnson is reputed to have said. A 'proof' that rocks are ideas would be a *reductio* of at least one of its premises (though which of its premises it refutes might easily turn out to be a question of genuine philosophical interest.)

What are your views on Connectionism and what place do they have (if any) in your theories?

I think the primary interest of Connectionism is that it has shown how hard it is for psychologists, neuroscientists and many in the AI community, to get it through their heads that Associationism isn't true. Associationism is the doctrine that learning complex consists of forming associative connections between its constituent concepts, and that thinking consists of the activation of such associations. So, to learn the concept BROWN SQUARE is to form an association between the concept BROWN and the concept SQUARE, and to learn the concept SQUARE is to form an 'associative connection between concepts *A* and *B* is one where there is a causal contingency between thinking one and thinking the other, the strength of which varies with the frequency that instances of one concept have regularly been followed by instances of the other in a certain creature's mind (it's in dispute whether there are other variables that also reliably affect the strength of associative connections; 'similarity' is often suggested, but for present purposes we can leave that moot). So, for example, for you to believe that cats have whiskers is for your mind to be so organized that thinking the concept CAT to reliably cause instances of thinking the concept WHISKERS. (Variants of this story available to accommodate psychologists of behavioristic bent, who don't talk of mental states, still less of causal relations among them. But here too we can skip the details). Given some version of association, the project of Connectionism is to simulate associationistic minds

on computers.

I suppose this may sound like a parody, but it isn't. There really are people, not only in psychology, but also in artificial intelligence, in brain science, and even in linguistics, who really do hold that minds are networks of associatively connected concepts, and that thinking consists of the activation such paths. After all, the brain, suitably enlarged, does look rather like it consists of many, many things (neurons) that are directly or indirectly, connected, one to one another by axons, dendrites and the like. If that's what brains are, then perhaps what happens in learning is that the activations of certain neuron comes to increase the probability that some of the neurons to which it is connected will become active too.

Earlier above, I deplored the regularity with which philosophers of mind have been content to proceed in ignorance of empirical findings about how minds work ('I am a philosopher, therefore do not bother me with data'). Connectionists are equally pig-headed only the other way around: 'I am a scientist, therefore do not bother me with arguments'. At least since the 18th century, the philosophical literature has offered what I take to be conclusive reasons for thinking that Associationism can't be true which, as far as I know, have never been seriously impugned. Thus, Kant argued that acquiring the belief that *As* are *Bs* can't be a matter of instances of *A* coming to cause instances of *B* because, Kant pointed out, because it is one thing to think lead is heavy but it is something quite else to think lead and then have that cause you to think heavy. The difference is that, in the former case, heaviness is predicated of lead, but in the second case it isn't. Accordingly, thoughts have truth values but associations don't. (Frege repeated this line of argument a century later, but psychology didn't pay any attention to him either).

The moral, I suppose, is that academic disciplines shouldn't be mistaken for 'natural kinds'. The fact that theory *A* is taught in the Department of Psychology but theory *B* is taught in the Department of Philosophy (or vice versa) is not a good reason for thinking that the truth of one is irrelevant to the truth of the other. Divisions between departments are there for the convenience of deans, not for the furtherance of research. The 'cognitive science' movement was, among other things, a self-conscious attempt to teach philosophers what they need to know about psychology and psychologists what they need to know about philosophy. It hasn't worked very well so far.

What was your motivation for writing LOT2? How have your views on LOT1 changed since you wrote LOT2?

This is a long story, and it's still unfolding. The latest version of the story is set forth in a new book with Zenon Pylyshyn (forthcoming from MIT Press in

2014) called *Minds Without Meanings*. The beginning was set forth in my *Language of Thought* (hereinafter ‘LOT1’ (1975)) a soft-cover is available version from Harvard University Press of is available for nostalgia freaks.

The basic idea was to revive the sort of ‘Representational’ theory of mind familiar from the British Empiricists (Hume was perhaps the prototype); but LOT1 wanted to make a couple of fundamental revisions. First, the Empiricists had thought that mental objects (beliefs, for example) are relations between minds and ‘mental representations’ (which Empiricists often called ‘Ideas’). LOT1 wanted to hold onto that: In the simplest kind of cases, to think about dogs is to entertain a mental object that refers to dogs (or, if you prefer, one that ‘expresses the concept’ DOG). However, though Empiricists had generally thought of Ideas as something like pictures, Berkeley offered very convincing reasons for saying that they can’t be; the question ‘What’s the alternative?’ remained moot.

According to LOT1, by contrast, mental representations comprise a system of structured objects not very different from the sentences of a ‘natural’ language (e.g. English). To think that’s a dog is to have in mind an object not unlike the English sentence “that’s a dog” (there are important reasons for denying that mental representations could actually be sentences of natural languages, but for present purposes, we can skip them.) So far, then, the LOT1 story was: We think in things that are more like sentences than they are like pictures. Concepts are the constituents of thoughts, in much the way that words are the constituents of sentences. The constituents of the thought dogs bark are the concepts DOG and BARK so arranged that the second is predicated of the first. OK so far.

The second major revision of Empiricism that LOT1 proposed was to abandon the Empiricists’ thesis that mental processes are associative. The suggestion was that, instead of thinking of learning as the formation of associations and thinking as their activation, LOT1 said we should think of them as computational processes defined over the structured mental representations that express thoughts. This was, in effect, to adopt the suggestion of Turing’s that eventually gave rise to computers. Associationism having proved a dead end, Turing provided an alternative.

The significance of Turing’s account of how the mind works (viz. by the manipulation of syntactically structured mental representations) really can’t be over-emphasized. For just one example, we’ve seen that if mental processes are supposed to be associations, there seems to be no way to distinguish two quite different processes that a mind might undergo: on one hand, forming the belief that lead is heavy; on the other, forming a causal chain in which in-

stances of the concept LEAD cause instances of the concept HEAVY. But if mental processes are computations – operations defined over linguistically structured objects, we can think of the belief that lead is heavy as predicating heaviness of lead, much as the sentence ‘lead is heavy’ does. In effect, Turing’s notion of mental processes as computations rather than association solved the problem that Kant and Frege had raised of distinguishing thoughts from mere causal sequences of concepts.

But if thoughts are structured arrays of concepts, where do the ideas come from? The Empiricists had an answer: All Ideas are associations of basic ideas, and basic ideas are traces of sensations. So the Empiricists’ psychology was thought to support an Empiricist epistemology in which all one’s concepts are definable in terms of sensations, and all one’s thoughts and theories are ultimately thoughts and theories about one’s sensations. Positivism followed.

One could easily imagine a computational psychology taking much the same path; and, in fact, almost all computational psychologists actually did, thereby raising a problem that still persists: Supposing that all concepts are definable in a vocabulary of basic (e.g. sensory) concepts; how is it possible to learn any new concepts? On the one hand, sensory concepts aren’t learned, they are just triggered by encounters with instances of the properties they express. Given the character of one’s sensory mechanisms, ‘learning’ the concept RED is just an automatic consequence of encountering red things; and, of course, the structure of the sensory mechanisms are innate (for this sort of reason, many of the Empiricists took the innateness of basic concepts to be self-evident). And since, by assumption, all the nonbasic concepts are definable in terms of the basic ones, it seems that there can’t be such thing as a kind of learning that increases the expressive power of the concepts one is born with. You can’t ‘learn a new’ concept because, by assumption, every concept you can learn is equivalent, by definition, to some array of concepts that you already have. This conclusion was all the more embarrassing because several decades of inquiry, both in philosophy and in cognitive psychology, showed that Wittgenstein was right: practically no words have definitions, sensory or otherwise.

LOT1 was inclined to say ‘well, so be it all concepts are innate’; in effect, that’s what LOT1 did say. But that conclusion struck many philosophers and psychologists as preposterous on the face of it; from time to time it struck me that way too. However, it’s one thing to disapprove of an argument’s conclusion; it’s quite another to show that the argument is unsound. What to do?

LOT2 reviewed this state of affairs and tried to put a bandage on the paradox by recruiting the notion of an innate ‘attractor space’ that had been developed

in the Connectionist literature. In effect it offered a kind of concept Nativism that was at least slightly embarrassing than the thesis that all concepts are innate. But I'm now convinced that it was unsuccessful. I now think I know the way out; it's one of the things presented in the forthcoming book with Pylyshyn. But I don't think you are going to like it.

For the last hundred years or so, the tradition in theories of conceptual content has taken for granted that concepts have both 'senses' and 'extensions', with the former determining the latter. The picture is very intuitive: The concept DOG has a senses such that it applies (has in its extension) all and only dogs. If the concept DOG is definable, then maybe its sense is its definition; maybe 'domestic canine' would do? Likewise for every other concept that isn't sensory. Acceptance of some or other version of this 'sense and extension' account of conceptual content (and hence of word meaning) has been practically universal in psychology, philosophy and linguistics. Indeed, there are lots of people philosophy who think that Frege had provided an argument that shows that this account of content must be right (see any standard introductory text in semantics.)

However, the 'sense and extension' account of conceptual content leads, by the kinds of considerations mentioned above, to the disconcerting conclusion that all concepts are innate. *Minds Without Meanings* hopes to dispense with this paradox. Very roughly, it says that there aren't any such things as senses (that would explain why nobody has been able to explain what senses are; in particular, it's why definitions and the like are so hard to come by.) There are extensions, of course – it's quite true that DOG applies to dogs and to nothing else – but the extensions of concepts are determined not by senses but by causal relations between minds and the world. The claim is that this is sufficient apparatus for a theory of conceptual content that is adequate both to the purposes of a Naturalistic psychology and, as a by-product, to avoiding radical Nativism.

The suggestion is that 'meaning' is a useless concept; that we can do quite nicely without it, and that it has been the root cause of many, many failures in theories of mind and theories of language. It may well strike you as itself unbearably paradoxical; it may strike you that there isn't much to choose between 'all ideas are innate' and 'there is no such thing as meaning'. But do read the book.

The BJUP team is extremely grateful to Professor Fodor for taking the time to answer our questions in such detail.

Interview: Huw Price

University of Cambridge

How did you get into philosophy?

Well, as is true of quite a few philosophers, it was a little bit roundabout. As an undergraduate in the early 70's at ANU in Australia, I started thinking I wanted to be a physicist, in particular an astronomer. So I started off doing physics and pure and applied maths. For some reason I did philosophy one as an undergraduate, but then went on in my second year to do physics and maths and it was only at the end of my third year that I decided I wanted to give up the physics and concentrate on the maths that I had to find some way of filling up my degree in place of the physics units I would otherwise have done, and I found I could do a couple more philosophy units and make up a major that way. And it was at that stage, when I was a third year undergraduate, that I started getting interested in the subject but it still wasn't for another two or three years after that that I finally made the switch, I even started graduate work in maths at Oxford that would normally lead to a PhD or a DPhil but I left after the first year of doing an MSc and moved to Cambridge to do philosophy at that stage.

What area did you focus on when you moved to Cambridge?

A couple of things; the person I moved to Cambridge to work with was Hugh Mellor who was someone I met as an undergraduate in Australia when he was visiting ANU as a young lecturer from Cambridge. And so it seemed natural when I got to Cambridge to start thinking about the kinds of things he was interested in. So I started thinking about probability, but also about time, at that stage Hugh was writing his book *Real Time*. There's a sense in which almost everything I've done since then grows out of those two interests. I went on to write my thesis on probability defending a kind expressivist account of single case probability judgment, an account very different from Hughes. All the work I now characterised as expressivism



or pragmatism grew out of this kind of generalising and thinking about the more general issues which came up, arising out of that approach to a particular case, the case of probability judgment. So that's where my interests in pragmatism came from. And then on the other side I retained this interest in time, although I was always interested in doing it in relation to Physics, and so my interests in philosophy of time and philosophy of physics grew out from that side. I also do work on causation which is somewhere in the middle. Maybe that didn't develop until a bit later, probably that developed during the 80's when I was back in Sydney and talking to Peter Menzies a lot, he was a big influence at that point.

What do you think the role of philosophy is nowadays?

It's a question I find hard to answer in any general way. For myself I think some of the most interesting philosophical problems are problems that arise at the intersection between philosophy and other disciplines. Where, in a sense, philosophy is like a certain kind of method or a certain kind of focus and interests in the nature of the concepts which are specific to those disciplines. Something of that kind. So I suppose my interests in philosophy of physics could all be characterised in those ways. They arise from places in contemporary physics where the techniques and particular aptitudes that some philosophers have can, I think, be useful in understanding what the physics is telling us about the world. But I suppose that not everything I do would fall into that category. In the area of pragmatism there's a sense in which it's more internal. The problem space is more one that is created by philosophers of the past. So much of it is trying to suggest that some traditional ways in which philosophers have approached certain issues – what is causation for example – are in a sense misdirected because it's more fruitful to approach via certain questions; Why do creatures like us have a need for a concept of causation? What role does it play in our lives? So there it's more internal, more generated by a reaction against a certain way of thinking about causation within philosophy. Then again that's not a general characterisation, but it's a characterisation of two particular kinds of philosophy that I find interesting.

Do you think philosophy is quite apt to interface with some of the sciences? Would it take on the role of an 'overseer' in checking their work in a 'meta-science' kind of way?

I don't think I see philosophy as an overseer. I see it as something between an assistant and a kind of next destination. If you're coming from some particular discipline you're interested in the foundational issues within that discipline. It's very natural to appeal to some systems in philosophy or move in the direction of philosophy itself. So there's a sense in which it's on the same level,

but it's sort of out to one side or I suppose you could think of it as further back. From when I was an undergraduate I always very much liked that idea that you find in Quine – there's not a sort of 'first philosophy' which stands above this.

What do you think the role of metaphysics is in philosophy and in interfacing with other disciplines?

I think quite a lot of things that people do under the heading of 'Metaphysics', fall into the categories of the kinds of philosophy that I think would be better done in other ways. People think about questions like; what is causation? What are laws? What is the nature of something? – issues. At least a substantial class of those issues are unproductive and there are better ways of approaching the puzzles from which they arise. Often they arise because there seems to be some reason for thinking that the notion in question is difficult to make sense of and difficult to reconcile with the scientific world view. This is most obviously the case in the case of ethical judgment – what is the place of morality in the scientific world? Those are the kinds of cases where I think that focussing on questions like – Why do natural creatures like us have a need for concepts of this kind? Or a use for concepts of this kind? These questions are more productive than thinking about the 'metaphysical' issues. In a sense trying to deflate the metaphysical view and replace it with something else. Now it's probably fair to say that not everything that goes on under the heading of 'metaphysics' is of that kind. The thinking I do wearing my other hat – thinking about the nature of the quantum world, some people would certainly think of that as metaphysics. I don't know, but I think that's rather terminological. I'd rather regard it as something continuous with physics. They might come back and say that in calling it metaphysics they weren't denying that in any case.

I certainly think that the moves I take away from metaphysics aren't really in any sense trying to disagree with them and that's what Strawson called descriptive metaphysics – the mapping out of the space of concepts and understanding how things hang together. I don't think it's something I do, myself, very much. But it's not something I regard as in any sense a sort of disreputable or misguided activity. On the contrary I think it's probably very helpful as a precursor to the anthropological issues. We need to understand the relationship between various systems of concepts in order to help us think about the question of the genealogy of where and why those concepts came into use.

Would you disagree with the role of meta-ethics in trying to find a metaphysical underpinning to notions like 'the good'?

I do want to disagree with that program, but I don't see that as disagreeing

with meta-ethics. It's just doing meta-ethics in a different way. Sort of starting the inquest again and thinking about the role of moral judgment and moral concepts. That's still doing meta-ethics in that it's in some sense stepping back from ethics. Not doing first-order ethics. It certainly involves a certain amount of philosophical distance, but it's not using that distance to think about metaphysical questions, it's using that distance to think about Anthropological questions.

What advice would you give to undergraduates? Both those that aspire to be philosophers and those who are just trying to get through their undergraduate degrees?

To people in the second category my advice would be to try as hard as you can to find the bits of the subject that interest you. It's a very diverse subject and of course depending on where you are and where you study it's not all going to be available to you. In my own experience when I first came across it there were just certain little bits that I found really engaging and interesting. And that certainly wasn't true, by any means, for all the bits of philosophy I encountered. This is probably good advice for many fields but I think it's particularly true of philosophy. To people who are thinking about pursuing further study – possibly a career in philosophy – I think again, thinking about the bits that interest you is very important, but also thinking about people and thinking about where you might find interesting people, interesting teachers, at the next stage, when moving to graduate study. Most people who go onto graduate work do some kind of research masters, so I'd encourage them to look around at the various research masters programs that are available in the UK and US and other places. Keep in mind that philosophy is a big field. Even within what the English speaking world is considered strict philosophy there are lots of people doing things in very different ways. Doing a little bit of research can make a huge difference to your graduate experience.

You most recently published a book called *Expressivism, Pragmatism and Representationalism*; would you mind giving us an overview of what it is about and the area?

Maybe a good starting point for that in the sense that it is something that a significant number of undergraduates would have encountered, especially if they've done some meta-ethics, they would have encountered the distinction between (and there are different names for this distinction but the two common names are) cognitivism and non-cognitivism. The non-cognitivists, the early ones, were people who were then called emotivists – people like A.J. Ayer – these were people who thought that when we make moral claims we're not making genuine statements about the world we are doing something else,

we are expressing our state of approval or disapproval or something like that. A more recent term for that kind of position about ethical claims, and other kinds of normative claims, is expressivism. That's associated with the idea that we're expressing some kind of state that, at least on the surface, isn't to be thought of as a belief about the world, it's be thought of in some other way. If you have a view of that kind then you have a picture where there's a kind of boundary between the claims which should be treated that way, like the moral and perhaps the aesthetic ones, and claims that should be treated as regular, factual claims. It's a very interesting question what goes into each camp, as it were, each category. People who do meta-ethics courses may not come across the fact that there are all sorts of other topics where philosophers have found the non-cognitivist approach attractive as well. Some of them have nothing to do with ethics. People who hold that sort of view about conditionals, about laws, about probability judgments, about causal judgments – a very wide range of things. And indeed about truth itself, this is what's now a very popular view about truth that when we say P is true we're not, as it were, talking about some property of truth. The role of the predicate 'is true' is not to be seen as ascribing some real, substantial property but in some other way. So then it has at least a family resemblance to the views about ethics that we were talking about a moment ago.

So there's that bunch of issues and in most views of this kind you'll find what Rorty calls a 'bifurcation' between the genuinely factual cognitive language on one side and the non-cognitivist stuff on the other side. What I guess you could characterise my view as is as somebody who thinks that when you kind of push on that there's an important sense in which nothing ends up on the cognitive side. So there's really no interesting work or a theory of language which is theorising about language in the way that people took for granted that you would theorise about language on that cognitive side. So notions like really talking about the world or really representing the world or really being factual, all those notions turn out to have not very interesting theoretical content. And one of the reasons for thinking that is closely linked to the view about truth I've mentioned a moment ago. Truth isn't a substantial property and that naturally goes with thinking that this idea that there's this relation between words and the world that turns out to not be an interesting way of theorising about language either. So there's a number of labels for the resulting position. I tend to call it global expressivism. It's also called non-representationalism or global pragmatism.

In some sense this is a view I've been attracted to for a long time and in my first book, *Facts and the Function of Truth*, I was at least moving in that direction. But I think I've found a better way of saying things over the years and so

the one way of characterising this view in Expressivism, pragmatism and representationalism is offering a defence of this global pragmatist view. The core part of it are some lectures I gave in the Netherlands a few years ago. There are three lectures and then the book also has responses to those lectures by four people – Simon Blackburn and Robert Brandom and Paul Horwich and Michael Williams who in various ways are quite sympathetic to some of these moves, although people closer to you are often the harshest critics. And then there's also a long section in the end where I respond to those essays and add some new material. Some of it comparing my views to those of Wilfrid Sellars and that's, at least from my point of view, a very interesting connection in that it's something I wasn't aware of until quite recently.

One other thing I could say, just as a point of reference, is that some people might be familiar with Simon Blackburn's work and his program called Quasi-Realism. Another way of thinking about what I am doing is as a kind of generalised Quasi-Realism.

You are also well known for your views on time – would you mind telling us a bit about time, the arrow of time and how philosophy can interact with physics?

One way into that would be to think about the way in which I oppose or stand opposed to another view point about time. One of the debates that people will be familiar with, if they've done any sort of course in philosophy of time or time within a metaphysics course at all, they'll be familiar with a debate between two sides called A-theorists and B-theorists, and these are terms that go back to McTaggart. Roughly speaking, the A theorists are the people who believe that there is an objective distinction between past, present and future. B-theorists are people think that that distinction is in a sense just anthropocentric. They think that 'now' is just like 'here' we use it to mark where we stand but it is not a part of reality. What's in reality is the distinction between earlier and later. One way of thinking about my position initially is as opposing the A theory and defending the B theory. Defending what is often called the Block universe picture – in which all time is briefly real as against a picture in which the present moment is picked out in some way, either as the only moment that exists or as, you sometimes get it, as a kind leading edge of a growing reality, the picture called the growing block view. We think that the past exists and the past is continually expanding as the 'now' moves through time. So there are various versions of these A theory views and I oppose all of them I think they are all more or less incoherent. In some of my recent work one of the things I've done is to look at what could be meant by the idea that there's a 'flow of time' which is another thing I'm kind of opposed to. And I distinguish three elements of this kind of package that tend to be not properly

distinguished. One is the idea that there's a distinguished present moment, which I mentioned before. The second is the idea that there is an objective direction to time, an objective distinction between past and future. The third is the idea that there is something dynamic, some notion of a passage, which is essential to time. I think all of those notions turn out to be either incoherent or nearly so. Perhaps the one that has the best prospect of being made sense of is that there is a direction of time but even that, when you really look into it, it's very hard to see what could be meant by that claim. The picture you end up with actually, if you accept those arguments, is not McTaggart's B-series, which would take for granted that there is an objective earlier/later than relation. It's what McTaggart called the C-series. Where you have the line, but the line doesn't have an objective direction on it, the relation is just between us. So you can say Kennedy's assassination was between the second world war and the election of Margaret Thatcher but you cannot say which one of those events came first. It was either war, assassination, Thatcher or Thatcher, assassination, war. But in reality neither of those depicts the right way. And this is one of the places I think it connects with Physics, because in physics the idea that there might be no objective direction of time came on the radar very precisely in the late 19th century in the work of Ludwig Boltzmann who's famous for work on thermodynamics and statistical mechanics. And what Boltzmann proposed was that the past and future are much like up and down, so if you live in the region of the universe where there is a thermo-dynamic gradient where the universe has been in low entropy and is approaching equilibrium, then in that region it would be natural to distinguish between the two directions. So creatures who live in that region would think of the direction in which they remember things as being the lowered direction and they would have expectations in the other direction. But Boltzmann suggests that's just like creatures on one side of the earth as seeing up in one direction and creatures on the other side of the earth seeing up in the other direction. But from the universe's point of view there's no real distinction between up and down.

That's one of the places where those rather metaphysical issues about time actually connect to the physics. Those are live debates because in physics, although it's probably true that most physicists believe in something like the block universe view, there are still examples around of very interesting and clever physicists who think there's something missing from that picture. There's a recent book by the physicist Lee Smolin, a very talented physicist but also a great writer, and he has this book defending the idea that there's something missing from the block universe picture and progress in physics depends on putting it back in.

That's my work on time, I guess, as it stands at the moment. The other thing

I'm interested in in philosophy of physics is the foundations of quantum theory and the question of what quantum theory is actually telling us about the world. There is this interesting possibility that hasn't been given the attention it deserves, which is that the answer to some of the quantum puzzles is that there's an element of retro-causality in the world, and I think that this is the most interesting sample of an area within contemporary physics where there is a deep roll for philosophical thinking, because I think there is an important possibility here, which may turn out to be the right possibility, the way the world actually is, which physicists are very, very ill equipped to think about, because it involves subtle thinking about bi-causation and questions about free will and agency, and questions about some of those questions that we were just talking about, about time. So then even the most insightful physicists who are interested in these issues about the meaning quantum theory, great, great thinkers like John Bell, haven't had the breadth of conceptual resources to tackle these issues in the way that they need to be tackled, in my view. So there's a real role here for bringing some clarity into this area.

One of my projects is working with a young physicist from San Jose, in California, a man called Ken Wharton who I started working with in the last three or four years and we now have a couple of papers. We have a plan to write a book trying to present these ideas in a semi-popular form, to make them more accessible to ordinary people as well as physicists and philosophers.

Do you think that the way philosophy interacts with physics now is how philosophy will begin to interact with other sciences and disciplines in the future?

I hope so, and I think that in the case of physics there has been a lot of progress over the last couple of decades. I know people in physics who were around before I was who remember what people were like in the field in the 60's and early 70's and they say that back then there was almost no interaction between physicists and philosophers. And they comment on the way that from how the 80's onwards there was a great growth of interaction, people going to common conferences and so on. I know in some ways it's been quite easy for me, I've had a lot of opportunities to go to conferences and talk to physicists. Cambridge has a very good tradition in this as do several other places in the UK and Europe and elsewhere. Now it's a very rich scene. Even if it's also true, and I think it's fair to say it is also true, that within physics and philosophy you'll find people who are a little bit negative about this activity. Physicists sometimes seem a little bit uncomfortable because their colleagues use terms like philosophy or metaphysics in a pejorative way. In philosophy you find instances of people being dismissive of some of the work in the philosophy of physics on the grounds that it isn't really philosophy, it's just physicists playing

with numbers and stuff. There is a wonderful community at the intersection now, but to a slight degree, and maybe this is inevitable, I think it made itself a community by distancing itself a little bit from the two communities.

I know a lot less about the intersection between philosophy and other disciplines but it's my impression that there are fascinating things going on there too. I know people who are interested in the intersection between philosophy and psychology for example, and I think that's very interesting. Some of my colleagues in Cambridge have contacts with linguists and with economics. On the whole I'm very optimistic about the continuing potential for philosophy to play a role in that sort of way. Because it's just the character of theoretical work that it tends to throw up these kinds of philosophical issues at the foundational level. So if you start thinking about the fundamental concepts of the discipline in effect you're going to be doing philosophy. That's good news for philosophers, good news especially for undergraduate philosophers, because it means that there's a sort of never ending supply of interesting problems to work on.

That reminds me of another piece of advice. That would be to take whatever opportunities you have to learn something significant about some other field as well. For example some branch in science. Another field which is clearly thriving that I haven't mentioned yet is philosophy of biology. I have many colleagues here, and used to have many colleagues in Australia when I was there, who work in those fields. What was characteristic of people like that and in other cases we talked about is that they often know a lot about the other discipline. So I think good advice for an undergraduate is to take opportunities as they come along to you, it depends what kind of program you're in. In the UK people are often in single honours programs which makes it much more difficult, but if you do have opportunities to study something else as well then I'd recommend taking them, if your philosophical interests are in any sense inclined in that direction.

The *BJUP* team is extremely grateful to Professor Price for taking the time to answer our questions, and particularly for speaking at the BUPS Autumn Conference 2013.

Interview: Shalom Lappin

King's College, London

How did you get into philosophy?

As an undergraduate, I was interested in philosophy. I started out as a kid reading Descartes and Plato and I became interested in some of the continental stuff. I discovered I was really interested in Logic and moved closer and closer to that direction in the field of analytic philosophy in Toronto (Canada, which is where I'm from originally). Then through interest in language got into Linguistics (which I pursued in MA studies at Tel Aviv university), I split my time in the MA between logic and semantics. Then I started my PhD in the states at Brandeis, and there I was working on logic and language; but I continued to do a lot of work in linguistics. So one of my supervisors was a logician, a philosophical logician, Fred Sommers (he was my official supervisor), and another who was a straight linguist from MIT, Ray Jackendoff. So I continued to teach when I went back to Israel in philosophy, but increasingly moving into linguistics, until I just basically became a linguist. When I went back to Canada for an extended period of time I was chairing a linguistics department, by that point I had become a full-fledged linguist because I just wanted to know the empirical properties of natural language rather than just speculate at them.

I then went back again to Israel and travelled a lot, and in the course of the wondering also found myself getting into computational linguistics, working in machine translation; specifically, English to Hebrew, at an IBM science centre in Haifa, and that was mostly because I needed money. Then I got very interested in computing, computing designs and applications of computational models to language as ways of understanding, which were more successful than even some of the theoretical models that linguists were using. So I continued to do logic and semantics, but got increasingly interested in computer science, went into computational research full time at IBM,



became a computational linguist, and that's more or less where I've been the whole time.

So it was following a problem essentially – what are the formal properties of natural language that allow us to learn it, interpret it, generate it, use it efficiently in the way that we do? And exploring it from the perspective of philosophy, logic, philosophy of language, linguistics, theoretical linguistics and then computer science and computation. I continue to work basically among those fields and, much to my pleasure, discovered that there is an area where this is very common for people not to worry about departmental boundaries, but just follow the problem and that's called 'Cognitive Science'. I would probably situate myself there these days. As a philosopher *cum* computational linguist, logician, and computer scientist working in cognitive science. I'm not too concerned about departments or disciplines. Fortunately the department at King's is very open and people enjoy interdisciplinary work. I was in the computer science department here for five years before I came back.

Do you see philosophy, and academic work in general, becoming less concerned with disciplinary boundaries and becoming more about the pursuit of knowledge?

Well you said "becoming", I think it traditionally was less concerned – if you look at Newton's chair – it was in 'Natural Philosophy'. Russell was doing the foundations of mathematics and logic; Frege and all sorts of people had science and maths backgrounds and then went into Philosophy, and people move back and forth all the time. Physicists from Einstein through to the founders of quantum mechanics were moving back and forth between philosophy and science. Turing, who created the basis for modern computing, was a mathematician who then got into the question of how to model computable functions. Psychologists have won Nobel prizes for decision theory work that's applied in economics.

What we are entering now, the more inter-disciplinary approach, where people find they need tools from other disciplines in order to pursue their puzzles, is the way science and knowledge was traditionally applied. Then it went through a period of rapid over specialisation, where people became illiterate in everything but their own field. That happened with a vengeance in philosophy. Whereas philosophy used to be the ultimate inter-disciplinary homeland for people, and took interest in anything, and you needed to know maths and science in order to pursue philosophy, it became a refuge from the facts and from mathematical sophistication. I think that's changing fortunately; as people realise knowledge is indivisible. When you're working on problems like the nature of knowledge, you can't afford to ignore the evidence from neuro-

science; and when you're working on the nature of physical reality, you cannot ignore what's going on in physics; and if you want to know about language, then you have to be interested in what the psychologists and the computer scientists are saying – and they are interested in what philosophy is saying too. I think that's a natural state of affairs.

Do you think that there's a tension between the 'deep' knowledge required for some areas and your view that knowledge should be pursued in a wider scope? Do you think the narrower view is forgivable?

It's understandable, but I don't think it's forgivable. I think the English education system in general is far too overspecialised. The Scottish system is actually much better; my youngest daughter did her undergraduate degree in Scotland, which was a four year BA course, and she was thrilled. Her degree was in economics but in her first two years she took a range of courses across different subjects. I mean here in a three year BA you rarely encounter materials out of your subject, unless you're doing a joint honours. This starts in A-levels, in that you're generally restricted to three A-levels. It's far too early for people to be dropping English and Maths. They're either illiterate about literature, or they're illiterate about set theory, or calculus and functions, or they're illiterate about physics. After thirteen years of school, the fact that they should come out skilled only in exam writing in particular areas is a waste of talent – people don't know what they want to do at 18. It's harder to learn these things when you're older, even though you're often more serious about them.

On the other hand, there is always the danger of becoming a dilettante. There are certain American-type programs that emphasise great ideas, history of great ideas, but without getting into any depth. You need broad based education that covers a whole range of areas where you are exposed, non-trivially, to areas of knowledge that you will need (or may need) and then allowed to choose a problem. The discipline itself is less important; what's more important are the problems that are guiding you. You can look at a problem from various points of view. It's always more helpful to look at it from a whole variety of disciplines - sciences and philosophy/arts. Very often when people are doing research in research labs, they will come together from a different set of backgrounds and try different methods for solving things. And they'll get stuck until someone from a totally different background comes and suggests a different perspective, because that's the way they looked at things. You ignore the insight of work going on in other areas and fields that are related to yours at your own peril, and your own disadvantage. There's a big tendency in philosophy to say – “well, that's just empirical”. Well yes, but if your theory of mind conflicts with well-established scientific findings then how viable is it?

Do you think meta-philosophical inquiry about the way we do philosophy and how philosophy should interface with other disciplines is important?

The problem I have with metaphysics is: what's wrong with physics? I mean I think people should do this naturally. Philosophy always was (and this is one of the things that attracted me to philosophy from earliest time and why I did all my degrees in it) the mother of all structured scientific rational inquiry. When a field became mature, it would tend to break off as a science, but it continued to remain in discourse with philosophy. If you look at the 17th, 18th, and even into the 19th century, philosophers were completely literate across a range of subjects and participated as citizens of those disciplines. The idea of departments and disciplines separating people to the point where they didn't have to look at each-others work is just not true. I mean there are explosions in research, and you can't read everything, but I think it is possible to be aware.

However, one of the advantages of digital technology is that you have accessibility, without having to ferret things out in libraries; you have access basically to everything going on. You also (and Britain is a real centre for this) have an intelligent, literate, popular presentation of scientific results. Mathematicians, physicists, psychologists, and a whole variety of other people can speak literally about the their work, in a way that's accessible to a general public; an intelligent (meaning: critical and committed) public that wants to know. Some of the documentary work on science that goes on in Britain is superb – it's stunning. Obviously, there is also bad science journalism, people looking for scoops and producing garbage, but much less so in extended documentaries on public television. It seems to me that there's less and less of an excuse for not keeping up. Though string theory is hard for instance (I don't know the maths), the general ideas can be acquired. If it's something that's important, you can go and talk to a physicist, and be direct with them, to achieve a level of insight into what you need to know. Similarly, if you're making claims about computational processes in the mind, and you don't know how an algorithm functions, or what a distributed model or a neural net is – you're going to talk either beyond the problem, or you're going to produce a certain amount of nonsense, and the best thing to do is to check it out. The information is there, and the ideas are there – you just have to decide that you don't need a passport in order to cross these boundaries. I had the privilege of taking a course with Quine when I was a graduate student in Boston, and he remarked that “disciplinary boundaries are very important to deans and librarians; they shouldn't be overestimated – the disciplinary boundaries that is.” His wonderfully crisp and ironic style says it all – why allow yourself to be stopped when you're investigating further? Why shouldn't philosophers who helped to create many

of these disciplines in the past continue to take advantage of them, and continue to contribute to them? There are also philosophers who have made very substantive contributions in areas, like: computer science, maths, physics, and psychology. My view is that I really don't care how people label themselves; I want to hear what they have to say.

We have another interview in the journal from Jerry Fodor; what do you think of computational theory of mind and Language of Thought?

When I first read *Modularity of Mind* by Jerry Fodor I was extremely impressed by it. I was quite young, it was the early 80's and I was also at least a fellow traveller in the whole Chomsky enterprise. I have since changed my views (fairly significantly). I now think Chomsky was asking the right questions, but giving consistently the wrong answers. Despite being interesting answers, worth exploring, and major contributions to the field, I think that whole 'nativist' approach to the language faculty as it's called (which encapsulated linguistics and cognitive modules) just looks, to me, to be misguided in light of a great deal of work in the last 15–20 years. With the interleaving and integration of cognitive processes from the ground up, this notion of encapsulation just doesn't look viable. Distribution of functions through the brain, plasticity of the brain – sure there's localisation effects, I'm not a neuroscientist but I'm reading broadly the same things they do, and it just doesn't seem like the mind comes clearly partitioned, so genetically determined cognitive packages just do not look right. Their theory of learning seems to me like a vacuous one, since everything's innate (except for a few parameters). I think it's been shown to be in some places confused, and in other places wrong albeit interesting.

So there's a lot more that's possible in the way of generalised inference across fields, that seems to be going on from the lowest levels of perception, from vision, right through to the higher cognitive function, and it seems to be playing a role in grammar induction and grammar acquisition too. So this is why I think linguists need to talk to cognitive scientists and computational learning theorists, areas that I've gotten into. Particularly computational learning theory – machine learning applied to actual modelling human learning, which suggests there's a lot more learning of a weak biased kind. That is to say, there are probably certain initialisation conditions for certain domains that are biologically determined, but the idea that there is something like a language gene doesn't make a great deal of sense. Geneticists like Stein and Fisher at Oxford, who study the FOXP2 gene, have shown that the attempt to correlate specific cognitive high level functions, like language acquisition, in use with the individual genes or gene sequences really doesn't work – it's far too complex to be reduced along this way. So I'm definitely not of the Fodor persuasion these days; I think he himself has started to change his mind on some of these issues.

This sort of radical innatist view that all concepts can be predetermined like anti-bodies, seems to be a bit of a *reductio* to me. Chomsky's syntactic theory has changed, but I don't think his basic views about the nature of language faculties have, and I have written extensively as to why it simply seems to not be correct; but that's an on-going, interesting argument with many sides to it.

I'd like to ask you specifically about a talk you delivered this year to the KCL department at Cumberland Lodge, about whether universities are public or private goods. Would you mind giving an overview for our readers?

The basic thesis which I was trying to advance is that it's possible to characterise higher education, like education generally, as a quasi-public good. By this I mean that it's the kind of public benefit that will not be efficiently provided by the market, in such a way at least as to assure universal access of high quality. That's a generally accepted view at the lower level of education through to secondary system, though not of universities. I tried to first of all give arguments as to why you would want to call it a public good, in the same way that the health service is a public good. With most people you don't have to argue with them here, I mean in the States you do, but in the rest of the Western world you don't. You do need to treat public health as a quasi-public good; as far as public health goes, you need a health system that provides high quality health care in a universal way. It needs to be publicly funded, possibly through taxes and contributors, but not in such a way as to deter access to care. Most people understand the connection between such a health system and the general well-being (economic and social) of a society. They do understand the idea of a high-school system.

There are of course people who would argue for market experiments, which generally produce the unfortunate results we've seen here and in the States, versus a kind of private school system that is a conduit for the wealthy to preserve privilege which is what it is in Britain. 7% of the population go to private schools, and 55% of the intake at Oxford and Cambridge comes from those sources, so there you have a system for preserving class. The idea that you need a universally funded high school system, which just comes out of public taxes and the basis for educating your population to a minimal level, achieving social mobility, has never been seriously controversial in any major country. Still for some reason, universities now are no longer treated, at least in government policy, as even a quasi-public good. The evidence for that is the evidence I provided in my talk, is from OECD (Organisation for Economic Co-operation and Development) figures – Britain has privatised its university system much more radically, and at a much faster pace, than any other country in the OECD. That includes Chile, which is the only country

that exceeds Britain in the proportion of private versus public funding in the system. Britain now has as of 2010 – somewhere around 25% or less of its educational system funded publicly, and the remainder is private. Most of that private funding is posed directly on households through the system of loans.

This is directly a result of government policy, and I tried to trace the policy. It is often seen as the revolution of the Tory imposed Browne report, but it very clearly started with Thatcher and ran through Labour. What Thatcher did was greatly expand the system, but not fund it. She didn't privatise it – she just created a *de facto* cap through underfunding, which meant the system couldn't continue to absorb people without crashing financially, because there was no tuition charge. This created a long term chronic funding problem that was generally handled on an *ad hoc* basis. Institutions were running deficits and governments covered them up to a point, and then vastly increased the intake of foreign students from outside the EU, who were charged very high tuition. The system began to breakdown, I think, in the 90's when the Labour government began, under Dearing, to charge tuition and privatise grants as loans. These were stop-gap measures. Tuition reached £3,200. It was actually Labour who commissioned the Browne report, and Browne was the former Chief Executive of BP, a company with not exactly a sterling record in a variety of areas. It was odd to appoint someone like that to overview the general direction of higher education (both funding and architecture). The report basically called for a publicly regulated (albeit entirely privatised) system with limited funding for market failure, with expenses produced by STEM subjects to keep them alive. This is because tuition wouldn't be sufficient if it was equal across disciplines.

The current government more or less accepted this call, although the liberal democrats insisted on capping at £9,000; but clearly the current system is not adequate to cover the cost. As things develop, I think you will see that, as most people who study the problem agree, the government basically produced a completely unworkable system. They are burdening students with very very high loans. The average loan burden for a BA is actually higher than in the States. There are some very hefty loans there for certain high earning degrees like law and medicine, but most people's tuition fees are lower. Moreover, interesting enough, there's public investment in the States, which, whilst being consistently lower than the OECD average, has been constant over fifteen years at around 35 percent. This represents heavy investment in state universities, as well as public investment in private institutions. Whereas here, we have gone from a publicly financed system of 80% public investment at the beginning of tuition fees in 1995, with 20% private, to about 25%, which is probably less now with the full effects of the Browne system, with around

75% being privately funded through loans. However, they remain public institutions under very rigid bureaucratic regulation. So they are run essentially as public business where students are consumers, and the revenue that sustains the institution are tuition fees. Notice these are undergraduate fees; there's no loan system in effect for postgraduate students, and there doesn't seem to be a plan to put one into effect.

The loan system puts the upfront costs of running the universities on the government, so while they claim that these loans are assets, because they'll be repaid, they will not be paid until fairly far in the future. Furthermore, there will be a high default rate due to people leaving or simply not being in a position to pay. Repayment starts at an early point at £21,000, which is a low salary, and to have such loans cripples one if you have a family. Graduating with upward of 40, 50, 60 thousand pounds in loans, if you include the maintenance, then having to take out mortgages etc. on top of that – the effects are obvious. The attempt to present it as a tax is false. Tax has nothing to do with earning power – tax (if it's progressive) gets higher only in proportion to your income over a certain threshold; but here the threshold is £21,000 so it's a flat tax, and it's very high. Plus the interest rates are subsidised by the government – it's clear that they want out of that, and that they're not going to be able to afford upfront costs. So there's something called 'the Rothschild report', that's been discussed in the press, which recommends privatising the loans by turning them over to banks. To the extent that that is successful, it will turn them into fully commercial instruments, and they will be discriminatory in excluding students who can't prove credit rates, or saddling them with very high, punishing interest rates. Postgraduate education (which is exactly what you want to encourage) is here now becoming increasingly unattainable, out of reach, for ordinary middle class families. Whilst it is largely free in the States, even at private institutions.

Basically, the view then is that education is a private good, and should be paid for by the consumer. Universities now are businesses that are run by a managerial class. All of this, by the way, happened under so-called 'New Labour'. I think I've mentioned that I regard New Labour as the ultimate intensional construction. Normally, if you have an extensional construction like "green house", then it's green and a house. With an intensional construction, like "Toy Car", it's not a car but it is a toy. "New Labour" on the other hand is neither new nor labour. So the current system I think basically represents a precipitous crash in the system. Under Thatcher and then successive Labour governments, the system was in constant crisis but it wasn't crashing. It was constantly being bailed out, struggling on, undergraduate admission was capped it wasn't expanding as quickly as it should have. There was con-

stant pressure to take more foreign students, and a discriminatory arrangement where foreign students had more or less priority since, because they paid full fees, they were not capped. So you had home students of equal ability who couldn't get in – the cap had already been reached. Now what you have I think is a system that simply will not work at all. The indications are that, already, like we're seeing here at King's, there's a huge pressure to accept large numbers of undergraduates – with our targets being increased on a regular basis. I think that's true across the sector – King's is by no means unique. Concomitant with that is the tendency to lower standards; so missed grades, and offers with lower grades, are becoming the norm in a certain percentage, in order to sustain enrolment, as it is the main source of revenue. There is also a danger of revenue generated by the humanities and social sciences being used to subsidise more expensive laboratory sciences like medical sciences, and again I think is not unique to King's – I think it's true all over. I don't think this is done out of malice on the part of management – I think this is the economic necessity of the system in which students have become the consumers who sustain the system – it's not a public good or a quasi-public good. On this basis, that's basically the only way the place supports itself. Quality then goes down – you get the minimum quality with the maximum sales.

This, I think, is catastrophic. The country is playing with the future of graduate education and research at a time when these are more vital than ever – for sustaining a technologically sophisticated, productive economy. Similarly, there is a danger of either massive decline in enrolment, or alternatively a tendency of institutions to offer low cost education that is of comparably low quality – relying largely on online teaching, where what you're doing is basically taking correspondence courses. It might be very useful, but it's not an education. It's not a full undergraduate education that gives you the basis for graduate work. All of these things are becoming acute and interestingly enough there's no public debate. The Labour party has nothing to say about this, having instigated the process and then allowed the Tories to complete it. The Liberal Democrats have traditionally promised free tuition, but (notice) with a graduate tax that is basically just a life-long loan system. So the idea that they were progressive on this is a misconception – they had a view not far from what's implemented. But they were happy to go along with capped tuition. I don't think the change is as radical as people make out. They were also basically arguing for a privatised system where it's the consumer who pays.

The argument for this system is, I think, a seriously bad argument. It is that the beneficiary of the good is the one who has to pay for it. By this reasoning then, people who walk on sidewalks should be the one who pay for it. People who use schools and the health service are the ones who should pay. For

any service the primary beneficiary will be the person who uses that service – the point is that if it is of such benefit that you want universal access to be insured, then can the market provide it? Well in some cases yes, the market can give you mobile phone service at a reasonably accessible, near universal, fairly low cost. So you don't want to turn mobile phone services into a state concession – as you probably would get worse service, if that's possible, and higher prices. This is however certainly not the case in education or in medical services and a wide range of other services – the market does a really bad job. Again, look at the British system. If you want high quality education, you need to go to a public or grammar school. The grammar school is publicly funded, but generally not accessible. Private schools are accessible, but just to people who pay. This is, I think, what's going to happen to the universities – you'll have a small elite that charge very high tuition and will attract very good students through reputation (deserved or not). Large numbers of people either can't afford to go to these universities or will go to very cut-rate places of questionable quality. And this is catastrophic, both for the British economy, and for its social and political future.

The fact that it's not an issue of public debate is something that should be worrying people – and the question is, why is it not an issue of public debate? I've been thinking about this for a long time, and I'm not sure, but one reason is that the staff (the academic faculty) are themselves not willing to discuss, it aside from a few hardy souls. Most people when you try and raise the issue will perhaps privately agree and complain about the system, but there's a great deal of inertia and fear. The system now is such that (unique, I think, among western university systems) British universities do not have academic tenure. For a long time, I think this was treated as a kind of gentlemen's agreement – you didn't have official tenure, as it was abolished by Thatcher in 1988, but you did have a permanent job. Events of the last five to ten years have shown that that notion is illusory. The rise of a pseudo-corporate management at universities that has essentially decided to turn revenue raising, branding and marketing into the primary concerns of the institution have, I think, set its priorities and have made it clear that academic tenure means nothing. Even in this *de facto* mode they are happy to dictate to departments and individuals what their research should be. This is done for what they regard as financial reasons, as well as other ones motivated by market considerations.

So people are worried. There is a claim that academic freedom is entrenched in the law – well that's actually not true. Yes, you can say whatever you want, but that doesn't ensure your presence at a university. Universities are now in a position where they can throw anyone out. They say: “well this stuff doesn't happen to fit our planning and recruiting priorities, we're moving to

another area”; and “this area is not attracting sufficient students and might be making money, but we could make more money by abolishing such and such a department or folding it in with somewhere else”; well that’s the sort of thing that goes on all the time. *I* was targeted – this goes on in many different places. It is generally at a junior level, as it becomes more noticeable with senior people, but it still happens. So the absence of real academic freedom in the sense that you are secure in your position to do, not only your research, but to express views on the nature of the institution you’re in and the direction of academic planning/development in the entire country is, I think, extremely precarious. This causes people to be worried – it’s hard to get jobs, and you hold onto the one you get, and that produces an atmosphere of fear. There are also academics who join management, discover it’s profitable, and collaborate with the whole thing. For them it’s a good deal – they make a lot of money. I’m sure they’re honourable, in the sense that they’re doing their job as best they can. Yet the structure of this kind of an institution, its notion of what management consists in, will produce a certain set of priorities and you have to behave according to them. That’s how you’re judged – by your job description; but it won’t produce results that we’re interested in. We need a different kind of institution; one where learning, accessibility of learning and freedom of inquiry are primary features of the place.

Students are the ones who have real power, because in the new system they are consumers, and so if there’s any kind of focussed critique of what’s going on they would be the ones who would be heard. On the other hand, they are trapped in the system in which there is no real environment where they can express views that anyone will take seriously. All the parties express the same approach – a tendency towards privatisation – every area of activity is a good in itself.

So what can undergraduates do about this?

I think they should begin talking about it in an organised and systematic way – using their skills as philosophers and scientists to analyse the problem in a rational way, and look at the facts and suggest alternatives. As I said I think we’re being fed a load of nonsense by apologists for the radical privatisation we see in universities, as well as policies that lay waste to large sections of the economy. The place to start I guess is always in insistent pointing out the problems which are clearly afflicting us here as a result of these policies, and suggesting obvious alternatives. One should try to get people to at least discuss these issues and then hopefully galvanise opinion, and demanding that political classes as they are distributed amongst parties pay attention.

Still, this just doesn’t seem to be going on. All of the influence is top down –

there's no serious discussion. There's nothing more powerful than good ideas, but they just don't seem to be coming about, or if people say things, they are very quickly ignored and dismissed. I find that strange – the students have most to lose here. They are the ones who face lifetimes of excessive debt and diminished economic prospects in a system where the people who are imposing this are the ones who benefited from free education, and are now insisting that they want to downsize and eliminate these services in the interest of economic interests that they seem to be entangled with – why accept that? Why allow established privilege to dictate social arrangements that are clearly not in the interest of a modern, forward looking, productive, egalitarian society – which is presumably what everybody wants? Why accept stratification, decreased social mobility and de-education of a population, technological decline and economic stagnation?

These are very practical problems, but the policies that are driving the current developments are presented as: required, necessary, and unavoidable. I just don't see that. There was what I called the free-rider problem. Large business and financial interests basically take a free ride on what was until now a public system and making sure that they lobbied against high taxes of any sort. Why do people accept these arguments? There's a great deal of economic research to show that there's no correlation, in any statistically interesting sense at all, between level of taxation, in a progressive tax system, and level of productivity, in fact countries that have cut corporate taxes and taxes on very high-income groups have not sustained a higher level of productivity than say Scandinavian countries which have a higher level of taxation and more developed social programs. But oddly enough those Scandinavian countries are beginning to very quickly emulate the British and American models.

It's a mystery to me – this is the country that pioneered progressive social reform in a democratic framework, avoiding violent revolution, and it just seems unable to sustain an interesting social and political discourse. Everybody got tired.

The *BJUP* team is extremely grateful to Professor Lappin for taking the time to answer our questions in such detail.

Issue 7(1)
Autumn Conference 2013

Properties, Predicates, and Evans' Argument against Vague Objects
Chris Blake-Turner

The Causal Theory of Reference, Sameness-Relations and Quasi-Constructivist Essentialism
Sam Carter

Verificationism and Relativism: A New Defence of Metaphysics
Atticus Gatsby

A Critique of Jeffrie Murphy on Resentment
Bjorn Wastvedt

Tackling the Repugnant Conclusion
Helen Weir

Socrates' Philosophical Mission in the Apology: Irony and Obedience
Corrado Musumeci

Environmental Philosophy and the Teleological Argument: A Polemic
Thomas E. Randall

On the Relation Between the Form of Beauty and Eudaimonia in Plato's Symposium
Carl-Otto Frietsch

Hegel's Philosophy of Right: Civil Society and International War
Joel Mason

Are Scientific Revolutions Irrational?
Sophie Osiecki

Interviews

Jerry Fodor

Huw Price

Shalom Lappin



Journal of the British Undergraduate Philosophy Society

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