

Issue 11 (2) Summer Conference 2017

*British
Journal of
Undergraduate
Philosophy*



Editor-in-chief: Ella Langham
London School of Economics

Journal of the British Undergraduate Philosophy Society

BRITISH JOURNAL OF UNDERGRADUATE
PHILOSOPHY

British Journal of Undergraduate Philosophy

The Journal of the British Undergraduate Philosophy Society

Issue 11 (2) Summer 2017

Editor-in-chief: Ella Langham

Co-editors: Nathan Oseroff
Bede Hager-Stuart

Manuscript editors: Conor Thompson-Clarke
Tommy Hewlett
Sophie Osiecki
Matei Gheorghiu
Benedetta Delfino
Oli Woolley
Shike Zhou

Commissioning editor: Edmund Smith
President of BUPS: Eugene Chua

British Journal of Undergraduate Philosophy

www.bups.org.uk

All rights reserved.

©2017 British Undergraduate Philosophy Society, and the authors.

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

Cover image used with the kind permission of Sophia Ben-Achour.

Editorial

There are times in which a group of people must act in great haste, worrying whether or not they will be able to construct some unwieldy conceptual ziggurat. Waiting until the eleventh hour to begin may be, for some, a way of moving against existence, feeling the pressure upon one's shoulders, like a nub of chalk drawn across a blackboard. Some coders even elevate the subject to a 'sport' of sorts, attempting to build a game in less than twenty-four hours; some academics make it their job, procrastinating professionally, churning out papers long after they are due. Of course, during those times, it almost always feels like flying by the seat of one's pants.

Thankfully, for almost the entirety of the time spent working on this volume (but not all!), *this* was not one of those times, and I believe we have now reached a comfortable stride.

I am fortunate to present another edition of the *British Journal of Undergraduate Philosophy*. Over this past year, it has been a pleasure to work with the now-*old* team on this labour of love: thank you to Eugene Chua, Victor Roulière and Edmund Smith for their tireless work and cooperation.

Sincere thanks is due to the team of manuscript editors: Tommy Hewlett, Harriet Ainscough Needham, Shike Zhou, Conor Thompson-Clarke, Matei Gheorghiu, Benedetta Delfino and Oli Woolley deserve our sincerest gratitude. A special thanks also to Bede Hager-Suart and Nathan Oseroff for devoting their time to formatting the journal in L^AT_EX, as well as doing final proofs and editing. A particular mention must be made of Nathan Oseroff for his last-minute work correcting unnoticed problems with references, bibliography, spelling, grammar and ensuring all papers adhered to house style. Furthermore, Sophie Osiecki deserves immense praise for tirelessly working in the position of Finance Officer. Without her, I can assure you, this conference would not have been possible.

Lastly, we would like to take this opportunity to thank Wolfson College at Cambridge University for allowing us the space to host the conference and to the University of Durham for their generous funding. We are also very excited to have Hasok Chang as our keynote speaker at the Summer Conference.

This edition includes papers ranging from egalitarian theory to Foucault and showcases some of the best undergraduate work in the United Kingdom. This volume features pieces we genuinely believe are worth sharing with the wider philosophical community, and hope they are received warmly. I hope you enjoy reading it as much as I have enjoyed working on it.

EL

Contents

- iii. Editorial

- 1. A Dilemma for global egalitarianism
Peter Wilson

- 9. Predicate Reference and Metalinguistic Nominalism
Beatriz Santos

- 18. Does Sarah Stroud's Account of Epistemic Partiality Conflict With Evidentialism?
Daniella Meehan

- 26. A Marxist Review of Justice
Giulia Carlini

- 32. Would a Life Without Death Be Intolerable?
Joshua Goh

- 43. Solutions to the Old Evidence Problem in Bayesian Confirmation Theory
Matthew Coates

- 54. Foucaultian Interpretation of Sexuality and Its Positive Implications: An Account of Change
Anna Giordano

62. A Kuhnian Critique of the Feminist Movement
Martha Crowe

A Dilemma for global egalitarianism^{*}

Peter Wilson

London School of Economics

Abstract

In ‘Cosmopolitan Justice and Equalizing Opportunities’, Simon Caney argues that equality of opportunity should extend globally because it is unfair if some people have worse opportunities because of their national identity. In this paper, I will suggest an interpretation of the requirements of global equality of opportunity (GEO) based on commensurate standards of living does not seem overly demanding. Drawing on one of John Rawls’ objections to global egalitarianism I will show, however, that a dilemma arises when one tries to universalise a principle of GEO. In our non-ideal world where countries with non-egalitarian values exist, trying to implement a principle of equality of opportunity globally may often lead to a dilemma of having to choose between actively coercing foreign states and not ensuring GEO’s future existence. Given that for a liberal egalitarian choosing to act on either horn of the dilemma has a high theoretical cost, with both options being unjust, I conclude that, for liberal egalitarians, GEO cannot be a universalisable principle.

Introduction

Global egalitarianism can be characterised as a family of views which hold that justice places limits on permissible global inequalities (Barry and Valentini 2009, 487). For a global egalitarian, international inequality is wrong not just because the global poor are in a state of absolute disadvantage, but because this inequality is in itself unjustifiable. On this basis, liberal egalitarian theorists like Simon Caney argue that equality of opportunity should exist globally. There are those like John Rawls, however, who oppose global egalitarianism despite affirming egalitarianism on the domestic level. These ‘moral statist’ typically argue that once every person is guaranteed certain minimum protections against starvation and poverty, the fact that some peoples are much wealthier than others does not undermine justice (Blake 2015). In order to argue for such a position, typically arguments are proposed which try so show that the considerations grounding egalitarian justice domestically do not also ground it globally (Barry and Valentini 2009, 488–489). In the following paper, I will take a different approach, arguing that global egalitarianism has a limited applicability, by showing

^{*}Delivered at the BJUPS Summer Conference, 8 July 2017 at Wolfson College, University of Cambridge.

that a dilemma arises when one tries to universalise a principle of global equality of opportunity (GEO). While I focus on extending Caney's arguments for GEO, I believe the dilemma poses a problem for many other proponents of global egalitarianism

For the purposes of this paper, I will assume, as Caney does, that principles of distributive justice apply at the domestic level due to the requirements of equality of opportunity. I will begin by explaining Caney's argument for GEO. I will then explain how GEO can be formulated so that it is plausible and not utopian; some such form of equal opportunity could hypothetically be provided by aiming for a commensurate standard of living, through equalising provisions of basic primary goods. I will then raise Rawls' objection to global egalitarianism on the basis of concern about undermining the liberal value of toleration, before showing how Barry and Valentini respond by showing how such concern might undermine domestic egalitarian demands on domestic non-egalitarians. This gives us reason to think there is no intrinsic reason to believe that not tolerating international political cultures is worse than not tolerating non-liberals in our own societies. Ultimately, however, I will argue that trying to create GEO in all circumstances leads to a dilemma of choosing between coercing foreign states and not ensuring the present and future existence of global equality of opportunity. Both forks of the dilemma lead to counterintuitive results, and therefore GEO is shown to be objectionable in some realistic non-ideal circumstances, discrediting the legitimacy of applying principles of egalitarian justice universally.

1 Caney's argument for global equality of opportunity

Caney argues that principles of distributive justice should apply at the global level because equality of opportunity should apply at the global level. Given that justice requires domestic equality of opportunity, he says that justice also requires GEO as they are supported by the same reasons (Caney 2001, 115). Equality of opportunity at the domestic level arises from the fact that people's life opportunities should not be based on arbitrary differences, like class status, or ethnicity, that arise due to being born to different backgrounds. Caney argues that a person's country of birth is as morally arbitrary as their class, status, or ethnicity, and therefore the life opportunities of people in different countries are arbitrarily unequal (*ibid.*). Therefore, Caney argues if justice requires equality of opportunity, it must also require GEO.

2 What GEO requires: commensurate standards of living

Understanding what exactly GEO requires is crucial to our discussion of its viability and legitimacy as a principle. An overly demanding interpretation of GEO might require that every person has perfectly identical opportunity sets. It would realistically

be impossible, however, for every person in the world to have equal opportunity to access a tennis court or equal opportunity to become a banker in Switzerland. A slightly weaker interpretation of GEO might suggest trying to use some calculus to account for the total value of every opportunity each person has and then equalise this value globally. This would also not be possible, though, as each person and each society will have different ideas on the value of access to different resources (Miller 2005, 62–63); in some societies access to tennis courts might be as important as access to churches, while in others it may be no more important than access to the cinema.

In answer to such concerns, Caney proposes an interpretation of GEO that does not require an absolute equality of opportunity, but rather, that the set of opportunities of two people living in different societies could be considered equal when they have commensurate standards of living (Caney 2001, 120). Caney's idea of commensurate living standards is grounded in the idea that a fair and neutral metric of opportunity can be constructed that relies on Amartya Sen's capabilities approach. The capabilities approach only assumes, fairly uncontroversially, that freedom to achieve well-being is of importance to all people, and that this freedom should be understood in terms of people's capabilities, which are their real opportunities to 'do and be what they have reason to value' (Robeyns 2016). Achieving this freedom requires important human goods, and Caney suggests Nussbaum has created a plausible list, which includes (but is not limited to): life, health, the avoidance of pain and independence. These are examples of opportunities that are important to all people and therefore could be equalised without a metric problem (Caney 2001, 121).

Caney's interpretation of GEO, therefore, would be a circumstance where 'people of equal talent have equal access to positions of an equal standard of living', where equal standards of living are determined to be those that equally contribute to one's well-being (ibid.). He does, however, admit that quantifying and measuring the quality of standards of living and their contribution to wellbeing, by relying on Nussbaum's list, is extremely difficult. Therefore, it would be nearly impossible to be sure that such global equality had been achieved, and therefore truly implementing it would be virtually impossible. In this way, Caney's interpretation could be criticised as being utopian.

I believe, however, that this criticism is unfounded. Sen himself helped create the Human Development Index which attempts to measure human wellbeing in a country by considering factors such as people's lifespan, education level and the GDP per capita (About Human Development). While this may not allow us to perfectly quantify the goods in Nussbaum's list and thereby fail to truly realise exact equality of these goods persons, it offers us an example of a neutral metric of standard of living. If we try to equalise a metric which even roughly approximates Nussbaum's measure of capabilities, such as HDI perhaps, we are almost certain to improve the state of global inequality worldwide. As Caney rightly points out, just because perfect capabilities equality may be impossible to measure or implement, this should not prevent us from setting

such equalisation as a goal and getting as close as we can to succeeding (Caney 2001, 122).

3 Rawls' objection to GEO and Barry and Valentini's response

Rawls rejects global egalitarianism on the basis that imposing global equality on the world as a whole would fail to show adequate respect for international cultural pluralism (Rawls 2001, 59–60). Rawls' concern is that imposing an egalitarian global order would involve disregard for those political cultures which differ from liberal ones but are decent enough to honour basic human rights (Barry 2009, 506). From this point onwards the term 'decent' will be used to describe countries or societies that at a minimum respect fundamental human rights and ensure a basic level of sufficiency for their people.

Even if GEO is a valid principle of justice, the value of not imposing on people values that they reject conflicts with GEO and is one liberal egalitarians must take into consideration (Caney 2001, 129). Liberalism as a doctrine is usually thought to be centrally committed to the value of toleration of different political cultures. Since domestically liberalism tolerates a variety of 'reasonable' conceptions of the good, on the global plane it would seem to follow that it should do the same, and this includes tolerating political cultures which are not fully liberal (Barry 2009, 507). The idea is that so long as societies respect fundamental human rights, are peaceful, and assist each other in time of absolute need, the world is just and societies cannot be objected to on the grounds of unreasonableness. Barry and Valentini point to a potential flaw in Rawls' argument. If it is reasonable for non-liberals internationally to not be governed by egalitarian principles, why it is not reasonable for non-liberals in a society with egalitarian beliefs to reject being ruled by egalitarian principles (ibid.)? After all, for those non-liberals born into egalitarian societies the only reason they must abide by these laws is due to the morally arbitrary fact of the location of their birth. Therefore, if we must tolerate decent political cultures to such an extent as to undermine equality of opportunity, it would seem to follow that this would also be the case domestically. Using such an account of reasonableness casts doubt on one's ability to justify egalitarianism at the domestic level.

If domestic equality of opportunity is a requirement of justice (as has been assumed for this purposes of this paper), however, and even domestic non-liberal are required to redistribute resources despite the value of toleration, Barry and Valentini suggest that there is no reason to believe that the value of toleration should undermine the duty of non-liberal internationally to redistribute. Rawls has not shown that there is an intrinsic reason to think that not tolerating international political cultures is worse than not tolerating non-liberals in our own societies.

Barry and Valentini are right to point out the potential inconsistency in an egalitarian defending domestic egalitarianism and global statism; it may well be the case that an egalitarian cannot justify statism in ideal circumstances, or in many circumstances at all. There are serious repercussions, however, if consistent egalitarians must believe that it is unreasonable for non-liberal societies internationally to reject egalitarian principles. While such repercussions would not exist in a world where a global state exists and in which there were no non-liberal societies, this is not the world we currently live in nor are likely to live in future. It seems to me that the existence of a world with distinct states, some of which are non-liberal, is a situation that any egalitarian principle has to be able to account for successfully. In what follows I aim to show that the high theoretical cost of imposing GEO in the many non-ideal circumstances possible in our world, indicates that GEO cannot be a universalisable principle.

4 A dilemma of toleration vs equality of opportunity

Currently, most of the world's richer nations are western liberal democracies. If these nations willingly transferred and created resources in poorer nations, this would be working towards GEO without any concerns about undermining the value of toleration. To ensure GEO for the future, however, we would need to get all countries to agree to redistribute to other countries poorer than them as it is possible that many decent non-egalitarian countries may eventually become the world's richer nations.

Imagine that despite all efforts to incentive it to do otherwise, such a decent country refuses to accept a resource transfer to help equalise opportunities, or chooses to accept resources at one point but refuses to help other countries through redistribution to other nations that require opportunity equalisation. While an egalitarian could argue that it is unreasonable for non-liberal societies internationally to reject egalitarian redistribution, the question remains as to what course of action should be taken in this circumstance where GEO is not complied with by these societies.

In this situation a dilemma arises: a proponent of GEO would have one of two options, both of which have a high theoretical cost. The first option would be to coerce the country into accepting or redistributing resources, thereby imposing egalitarianism on the society. A strict egalitarian who emphasises equality above all other values may be happy to take such action. Caney, along with most other global egalitarians, however, also commit himself to liberalism and takes seriously the value of toleration. When incentives have failed to make a country act in accordance with GEO, coercive methods would need to be used to force compliance. It would seem, however, that this would be unjustified: all semblance of tolerance of decent nonliberal societies and their values would be given up if we were to actively impose on them values they did not agree with through the use of force. We would unavoidably be sending them the message

that their value systems are wrong and that it is unjustifiable to hold the views that they do. In doing so, we would also be undermining the country's sovereignty and ability to self-determine the decent societal values they uphold.

The other option would be to accept that this country will do nothing to aid attempts to create GEO. Take an example where a poorer decent country refuses to accept a resource transfer to help equalise global opportunities: it would seem unfair to the state's residents for an egalitarian to simply accept that based on the morally arbitrary fact of birthplace, they will not benefit from the global scheme of GEO when there are certain to be those who would be glad despite the society's overarching values. In an example where a non-liberal country has previously received resources from willingly donating egalitarian states, but once it has gained wealth refuses to help other countries through redistribution to poorer countries it seems we are tolerating a country's values at the expense of fairness. It seems intuitively unfair that a country can benefit from a global redistributive scheme and not contribute to it. The only alternative, having earlier ruled out the use of significant coercive force, would be to deny a country access to the redistributive scheme in the first place on the basis that it lacks egalitarian values and likely will not contribute. But this also is unjustifiable to a liberal egalitarian. It would seem to unfairly treat those citizens of the country that would appreciate and benefit from a redistributive scheme as undeserving of the equality that the rest of the world has based on what their society potentially may not do to aid the scheme and on the morally arbitrary fact that they were born into a society which as a whole rejects egalitarianism.

In summary, the dilemma that GEO non-compliant countries create forces a proponent to either illiberally coerce a foreign nation into action (thereby violating toleration, sovereignty and self-determination) or to accept and allow for clear ingalitarian unfairness to occur. Choosing the latter horn is also likely to severely undermine GEO globally in circumstances where the world's richer countries are decent non-egalitarian states. Both horns of the dilemma seem to have consequences with a high theoretical cost.

5 Caney's possible response: a case-by-case approach?

Caney might attempt to respond to this dilemma by taking suggesting that in general in non-ideal circumstances where countries are non-compliant with global egalitarianism we must balance the liberal egalitarian values of tolerance and equality of opportunity. He believes that in some cases the latter should take precedent, as he states that there is 'no reason to think that 'cultural justice' should subordinate 'economic justice' (Caney 2001, 131). In some of the relevant cases, therefore, he might argue that choosing the first horn of the dilemma form and using slight coercion might be jus-

tified, while in other cases he might suggest choosing the second horn as in certain cases there is a greater cost of undermining toleration than undermining of present and future GEO.

What is clear, however, is that there are possible cases of the dilemma where it is unclear how there can be any balance. In particular consider the possibility that many of the world's richer nations ended up being non-liberal countries: GEO would be entirely undermined unless they were somehow forced to redistribute, with incentive methods unlikely to work in this circumstance. In many possible cases, either GEO is mostly given up, or any semblance of toleration is given up when a country is actively forced against its ethos to redistribute resources to another society that has fewer opportunities than it does. This in itself shows us that GEO as a principle is not universalisable as in some feasible circumstances working towards it even slightly would be seen to be unjust by many egalitarians as it would require entirely undermining the liberal value of toleration. It seems that GEO is not a principle that is just to apply in all situations, and it, in fact, may motivate seriously unjust action if complied with in certain circumstances.

6 Conclusion

In conclusion, I have shown that for liberal egalitarians like Caney, GEO is not a universalisable principle because ensuring its future for all time without the existence of a global state would be almost certain to threaten the liberal value of toleration severely at some point. It seems in our non-ideal world, equality of opportunity between liberal democratic societies who have a shared ethos of GEO could achieve such equality between themselves, but enforcing such a system globally for all time would be illegitimate due to the potential to undermine the liberal value of toleration entirely.

References

- [1] About Human Development — Measure of America: A Program of the Social Science Research Council Retrieved from <http://www.measureofamerica.org/human-development>
- [2] Christian, B. and L. Valentini. (2009). 'Egalitarian Challenges to Global Egalitarianism: A Critique'. *Review of International Studies*, 35(3): 485–512. doi:10.1017/S0260210509008626.
- [3] Blake, M. and P. T. Smith. (2015). 'International Distributive Justice.' In *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/spr2015/entries/international-justice/>.
- [4] Caney, S. (2001) 'Cosmopolitan Justice and Equalizing Opportunities.' *Metaphilosophy*, 32(1–2): 113–134. doi:10.1111/1467-9973.00178.
- [5] Miller, D. (2005). 'Against Global Egalitarianism'. *The Journal of Ethics*, 9(1–2): 55–79. doi:10.1007/s10892-004-3319-6.
- [6] Rawls, J. (2001). *The Law of Peoples*. New Ed edition. Harvard University Press.
- [7] Robeyns, I. (2016). 'The Capability Approach'. In *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.). <http://plato.stanford.edu/archives/win2016/entries/capability-approach/>.

Predicate Reference and Metalinguistic Nominalism

Beatriz Santos

University of Manchester

Introduction

In ‘On What There Is’, Quine (1948) says the following:

That the houses and roses and sunsets are all of them red may be taken as ultimate and irreducible, and it may be held that McX is no better off, in point of real explanatory power, for all the occult entities which he posits under such names as ‘redness’.

I take this quote to summarise Quine’s view that predicates are not referring expressions, i.e. they do not refer to universals (properties and relations). In this essay I argue predicates are not referring expressions. However, I claim that Quine’s nominalism is not ambitious enough. Instead, I argue metalinguistic nominalism can genuinely justify the claim that predicates are not referring expressions.

I begin by discussing Quine’s argument for his view (§1). I then introduce Strawson’s (1968) and Jackson’s (1977) arguments against Quine’s view (§2–3). Finally, I argue against Strawson and Jackson’s positions by introducing an argument for metalinguistic nominalism (§4). I conclude that Sellars’ (1963) metalinguistic nominalism can justify the claim that predicates are not referring expressions.

1 Quine’s disquotational theory of predication

Consider the subject–predicate sentence ‘Quine is smart’ (‘a is F’). ‘Quine’ is the subject term, a proper name. It picks out the object so named, Quine. ‘(Is) smart’ is the predicate term, which describes the object named. Predicates are thus conceived as what the rest of the sentence says about the object named.

Consider the context of Quine’s quote. Quine is concerned with what there is in the world. Assuming that our language describes an objective reality, then true subject–predicate sentences contain predicates that appropriately describe reality. What enables predicates to do so? Realists and nominalists answer differently, as follows:

On the one hand, realists claim that objects which are similar hold in common a universal. The nominalist, on the other hand, denies this: there are only particulars. What

problem gives rise to this dispute? Realists argue that we can only explain the phenomena of apparent similarity, predication, and abstract reference by positing universals. The realist claims that red houses and red roses are red because the universal *redness* holds of both of them; the truth of subject–predicate sentences can only be explained if the predicate term refers to a universal; and the truth of sentences including abstract singular terms (ASTs)¹ can only be explained if ASTs are names of universals (Lou 2006, 17). Quine thinks that nominalism can deal with these phenomena without postulating, like the realist (McX), universals. Before the previous quotation Quine writes:

One may admit that there are red houses, roses, and sunsets, but deny [...] that they have anything in common [...] the word ‘red’ [...] denotes each of sundry individual entities which are red houses, red roses, red sunsets; but there is not, in addition, any entity whatever, individual or otherwise, which is named by the word ‘redness’ (Quine 1948, 29–30).

Two related confusions about what is said here must be cleared away. First, even if the realist is wrong to posit universals (‘occult entities’) ‘under such names as ‘redness’’, it does not follow that the adjective ‘red’ is not a referring expression. Adjectives are not names. Second, even if the realist is wrong in thinking that predicates are a special kind of name (names of abstract entities), it is still possible that predicates are referring expressions which are not names (MacBride 2008, 428). I will only show that predicates are not referring expressions when they are a special kind of name – names of abstract entities. The latter possibility remains open, but I cannot address it here due to space constraints.

Once these confusions are set aside, we can understand Quine as saying that we do not need to explain the phenomena of similarity at all. The universal redness is not held in common by red houses, roses, and sunsets. All they have in common is that ‘red’ is true of them, i.e. they are all red. We should take this ‘as ultimate and irreducible’. That they are all similar in colour is just the way things are. To think, like the realist, that there is an extra explanatory task to be achieved by universals is wrong. Positing universals adds nothing in ‘real explanatory power’ because there is no such task to be completed.

Once we accept this, explaining the phenomena of predication is straightforward. What makes the sentence ‘the house is red’ true is that the house is red. Subject–predicate sentences are true because things in the world actually are the way those sentences say they are. ‘Quine is smart’ truly tells us how the world is because ‘Quine’ refers to an

¹ ASTs are expressions in our language that are taken to refer to abstract entities like properties and relations (universals).

object (Quine), and ‘smart’ is true of Quine. Thus, all there is to the function of the predicate ‘(is) red’ is to be true of red things. The whole contribution of predicates can be given by the following disquotational schema:

(P) ‘F’ is true of x iff x is F (MacBride 2008, 429)

where the same predicate is mentioned on the left-hand-side and used on the right. Now, (P) is not quite right. Truth is relative to a language. The word ‘red’ is only true of red things in English. Hence, we need to replace (P) by

(P_L) ‘F’ is true-in-L of x iff x is F.

Substituting ‘F’ for ‘(is) red’, we have:

(RED_L) ‘Red’ is true-in-L of x iff x is red²

where L is English.

The way in which the nominalist deals with the phenomenon of abstract reference will be discussed, and criticized, in §3 and §4. For now let us focus on predication. Showing that predication can be satisfactorily understood as (P_L) partly involves showing that the realist is wrong in constructing predicates as referring expressions. In §2 and §3 I will put forward the realist argument for this. In §4 I will show that the argument fails.

2 Stawson’s argument

Let us first introduce some terminology. Say that referential position is the position occupied by referring terms; and predicative position is the position occupied by predicates (Quine 1960).³ Now note the difference between predicate terms like ‘(is) red’ (F) and ASTs like ‘redness’ (F-ness). In subject–predicate sentences (‘a is F’), ‘F’ appears in predicative position; ‘a’ appears in referential position, and seems to be the only referring expression. In sentences of the form ‘F-ness is G’ an AST appears in referential position, a predicate ‘G’ appears in predicative position, and ‘F-ness’ seems to be the only referring expression.

Even though ‘F’ does not seem to be a referring expression in ‘a is F’, Strawson (1968, 109–110) argues that it actually is because ‘F-ness is G’, where ‘F-ness’ is a referring expression, presupposes that ‘F’ refers. That is, Strawson argues that the predicate ‘(is)

²For a discussion of objections about the alleged circularity of (P) see MacBride (2008, 429–32).

³It does not follow that predicates are not referring expressions because, if predicates do refer, they can take (implicit) referential position.

F' is an implicit referring expression because the AST 'F-ness', an explicit referring expression, presupposes that 'F' refers.

Consider the following argument:

(1) Quine is smart.

(a is F)

(2) Smartness is a characteristic Harvard University appreciates.

(F-ness is G)

Therefore,

(3) Harvard University will hire Quine.

What is the relation between (1) and (2) that yields (3) as a conclusion? Strawson (1968, 109) thinks that there is something explicit in (2) that is only implicit in (1). This is made explicit by the expression 'characteristic of'. This expression makes explicit that it was implied in (1) that smartness is a characteristic of Quine.

In (2) 'smartness' occupies referential position and so explicitly refers to a universal, the property of being smart. Since it was implied in (1) that smartness is a characteristic of Quine, the predicate '(is) smart' in (1) has implicit referential position. Implicit because this is only clear when we understand that (1) implies that smartness is a characteristic of Quine. That is, when we understand that the predicate '(is) smart' is actually an AST, the name ('smartness') of a universal. '(Is) smart' has implicit referential position because it is implied in the AST 'smartness', which in (2) has referential position. Hence, '(is) smart' is implicitly referential because 'smartness' is explicitly referential.

I see this argument as a conditional statement: if ASTs are referential expressions, then so are predicates. Thus, it has to be established that ASTs are referring expressions. Jackson (1977) gives an argument for this.

3 Jackson's argument

The Quinean nominalist usually deals with the phenomenon of abstract reference by claiming that statements containing ASTs, which are putatively about universals, can be paraphrased as statements about particulars. Take the sentence, 'smartness is a characteristic of Quine'. Here 'smartness' seems to refer to a universal. But this can be paraphrased away. The resulting sentence is 'Quine is smart', in which the only referring expression is 'Quine'. Hence, the Quinean nominalist argues against Strawson by claiming that sentences containing ASTs can be paraphrased as sentences about particulars.

However, Jackson (1977) claims that this is not possible in all cases. He argues that sentences like ‘red is a colour’ cannot be paraphrased as sentences that are not about colour properties.

The nominalist usually paraphrases (4) as (5).

(4) ‘Red is a colour’ [logical form (Cr)]

where ‘Red’ is an AST allegedly referring to a universal.

(5) ‘Necessarily, everything red is coloured’

$\Box(\forall x)(R_x \supset C_x)$

where there are no ASTs.

Jackson argues that if the paraphrase from (4) to (5) is a correct paraphrase, so should be the paraphrase from (6) to (7):

(6) ‘Necessarily, everything red is shaped and extended’

$\Box(\forall x)(R_x \supset (S_x \wedge E_x))$

where there are no ASTs. (Because for something to be red, it must be possible to see it. In order to see it, it must have macroscopic size. And if it has size, it necessarily has shape and extension.)

(7) ‘Red is a shape and an extension’

$(S_r \wedge E_r)$

where ‘red’ is an AST allegedly referring to a universal.

Let us check if the paraphrase from (6) to (7) is analogous to the one from (4) to (5). In (4) ‘red’ is an AST that appears to be the name of a universal (‘redness’) and ‘(is a) colour’ is a predicate that describes it as a colour property. In (5) ‘necessarily’ takes wide scope, ‘red’ is a predicate, and ‘(is) coloured’ is a predicate behaving as an adjective-like expression.

In (6) ‘necessarily’ takes wide scope, as in (5), ‘red’ is a predicate, as is ‘red’ in (5), and ‘(is) shaped and extended’ is a conjunction of predicates behaving as adjective-like expressions, as in (5). In (7), ‘red’ is an AST that appears to be a name of a universal, as in (4), and ‘(is a) shape and (an) extension’ is a conjunction of predicates that describe that universal as having properties of shape and extension, as in (4). Inspection of the logical form of each of these sentences shows that (7) has the same logical form as (4),

and (6) has the same logical form as (5).⁴ Thus (6)–(7) are analogous to (4)–(5).

Given that red is obviously not a shape and not an extension, (6) to (7) cannot possibly be a correct paraphrase. Therefore, by reductio, (4) to (5) is not a correct paraphrase either. Jackson concludes that (4) must say something about a universal.

4 Metalinguistic nominalism

Metalinguistic nominalism holds that only particulars exist, but rejects the Quinean analysis of abstract reference. Sentences incorporating ASTs are not disguised ways of making claims about particulars. Metalinguistic nominalism takes a more ambitious stand, and claims that sentences incorporating ASTs are actually disguised ways of making claims about the linguistic expressions we use to talk about nonlinguistic objects (Loux 2006, 63). When the paraphrase of (4) is constructed according to metalinguistic nominalism Jackson's argument fails. Given that Strawson's argument relied on Jackson's argument showing that ASTs are referring expressions, his argument fails too.

Let us first consider Carnap's (1959) metalinguistic nominalism. I will then put forward a more powerful form of metalinguistic nominalism – Wilfrid Sellars' (1963) approach – which is an extension of Carnap's approach notable because it is not subject to a number of criticisms that target Carnap's.

Carnap claims that sentences containing ASTs are 'material mode of speech' or 'quasi-syntactical' sentences. They are sentences formulated as though they are 'object sentences, sentences about nonlinguistic objects, i.e. universals (Carnap 1959, 285). However, on analysis, they are 'formal mode of speech' or 'syntactical' sentences, i.e. disguised ways of making metalinguistic claims about linguistic expressions (Loux 2006, 64).

Sentences like (4), that contain ASTs like 'redness', are quasi-syntactical. On analysis, however, (4) is to be paraphrased as

(4a) 'Red' is a colour predicate

which is a syntactical sentence that eliminates (4)'s putative reference to a colour property.

Sellars (1960) posited two objections to Carnap's theory. Firstly, on Carnap's theory sentences incorporating ASTs are only true relative to the language of the ASTs. When

⁴Or can be constructed as having the same logical form, since the conjunction is only true when both conjuncts are true.

speakers of English utter (4), what they are talking about (the English word ‘red’) is something different from what Portuguese speakers are talking about (the Portuguese word ‘vermelho’) when they utter ‘vermelho é uma cor’. This means that if (4a) were the correct paraphrase of (4), (4) would only be true relative to English, since (4a) is only true in English. But, surely, (4) is true relative to all languages. Therefore, (4a) cannot be the correct paraphrase of (4).

Secondly, Carnap’s nominalism does not completely avoid commitment to universals. Rather, it replaces nonlinguistic universals for linguistic ones. In place of redness we get predicatehood, the universal held in common by everything that has the property of being a predicate (Brandom 2015, 239-240). Carnap’s theory merely replaces non-linguistic universals – the property of being a colour – by linguistic universals – the property of being a predicate.

Sellars’ (1963) metalinguistic nominalism is a development of Carnap’s theory that is immune to these two objections. Sellars solves the second problem by introducing the notion of distributive singular terms (DST). According to Sellars, the term ‘red’, as it appears in (4a), is equivalent to the expression ‘the word “red”’, or simply ‘the “red”’. He calls this a DST. A DST is an expression of the form ‘the K’, where K is a common noun. This expression seems to be the name of the universal K, but it is a device for indicating that a general claim is being made about the individual Ks (Loux 2006, 66–67). Sellars denies that (4a) commits Carnap to the existence of linguistic universals because ‘red’ in (4a) is implicitly the DST ‘the ‘red’’. (4a) is a disguised way of making a claim about the tokens of the word type ‘red’. (4a) says that all the tokens of ‘red’, in the context of (4), are colour predicates.

Sellars solves the first problem by introducing the notion of dot quotation. In the usual use of quote marks (‘ ’) we place single quotes around a word from a language to create an expression that picks out tokens of that word in that language (ibid., 68). Our metalinguistic discourse becomes language bound. We can solve this problem by noting that different languages have expressions that are functionally equivalent. The English ‘red’ and the Portuguese ‘vermelho’ are functionally equivalent: in their respective languages, ‘they function in the same way as responses to perceptual input; they enter into the same inference patterns; and they play the same role in guiding behaviour’, i.e. they have the same linguistic role (ibid., 68). Sellars introduces dot quotation as a form of metalinguistic discourse that brings out this fact. In dot quotation, dots are placed around a word from a language to create a metalinguistic common noun true of all those expressions that are functionally equivalent to the dot-quoted term (Loux 2006, 68). Thus when we dot-quote the English ‘red’ we obtain the English term *red*, that is true of all the expressions which, in their own languages, are functionally equivalent to ‘red’ in English. Thus the English ‘red and the Portuguese ‘vermelho’ are both *red*s.

Using dot quotation we can give a full-blooded nominalist paraphrase of (4),

(4b) The •Red• is a colour predicate

where ‘the •red•’ is not a name of a linguistic universal, but a DST. Thus, (4b) is really a general claim about linguistic expressions understood as tokens. It is the claim that

(4c) •Red•s are colour predicates.

(4c) is not subject to the first objection. Even though ‘•Red•s’ is an English term, it refers to all those expressions which, in their own language, are functionally equivalent with the English ‘red’. Thus, the truth of (4c) is not relative to a specific language (Brandom 2015, 244). (4c) is also not subject to the second objection because ‘red’ is paraphrased as ‘the •red•’, a DST, which is not a name for a universal, but a device for making a general claim about linguistic expressions (as tokens), the •red•s.

This theory is not subject to Jackson’s criticisms either. The nominalist should not paraphrase (4) as (5), but rather as (4c). Thus, that paraphrasing (4) as (5) leads to an absurd conclusion – (7) – is not a problem for metalinguistic nominalism. The metalinguistic nominalist correctly paraphrases (4) as (4c). On this analysis, (4) is a disguised way of making a metalinguistic claim about linguistic expressions (understood as tokens) which are functionally equivalent across languages. Thus, Jackson is wrong in claiming that (4) says something about a universal. He fails to show that the AST, ‘red’, in (4) is a referential expression. Since Strawson showing that predicates are implicitly referring expressions depended on Jackson showing that ASTs are explicitly referring expressions, Strawson’s argument also fails. Predicates do not implicitly refer to universals, because ASTs do not explicitly do so. Predicates, when they are a special kind of name – names of abstract entities – are not referring expressions.

In conclusion, Jackson’s argument is a clue for why Quine’s nominalism is not ambitious enough. When trying to justify the claim that predicates are not referring expressions, it is not enough to say that sentences putatively about universals are actually sentences about particulars. If this were the case, Jackson, and consequently Strawson, would have been successful in showing that ASTs, and predicates, are referring expressions. However, when we adopt a more ambitious version of nominalism, such as Sellars’ metalinguistic nominalism, we can justify Quine’s claim that predicates are not referring expressions. We do this by claiming that sentences putatively about universals are actually disguised ways of making a metalinguistic claim about linguistic expressions (understood as tokens) which are functionally equivalent across languages. Even though this is not a *prima facie* intuitive reading of sentences like (4), its success in showing that Jackson’s and Strawson’s arguments fail is sufficient reason for both realists and nominalists to pay close attention to this alternative version of nominalism.

References

- [1] Brandom, R. B. (2015). *From Empiricism to Expressivism: Brandom Reads Sellars*. London: Harvard University Press.
- [2] Carnap, R. (1959). *The Logical Syntax of Language*. London: Routledge & Kegan Paul Ltd.
- [3] Jackson, F. (1977). 'Statements about Universals'. *Mind*, 86 (343): 427–429.
- [4] Loux, M. J. (2006). *Metaphysics: a contemporary introduction*. Oxford: Routledge.
- [5] MacBride, F. (2008). 'Predicate Reference', in E. Lepore & B. C. Smith (eds.), *The Oxford Handbook of Philosophy of Language*. Oxford: Oxford University Press: 422–474.
- [6] Quine, W. V. (1948). 'On What There Is', *The Review of Metaphysics*, 2(5): 21–38.
- [7] Quine, W. V. (1960). *Word and Object*. Massachusetts: The Massachusetts Institute of Technology.
- [8] Sellars, W. (1960). 'Grammar and Existence: A Preface to Ontology', *Mind*, 69(276): 499–533.
- [9] Sellars, W. (1963). 'Abstract Entities', *The Review of Metaphysics*, 16(4): 627–671.
- [10] Strawson, P. F. (1986). 'Singular Terms and Predication', *Synthese*, 19 (1/2): 97–117.

Does Sarah Stroud's Account of Epistemic Partiality Conflict With Evidentialism?*

Daniella Meehan
University of York

Introduction

The aim of this paper is to assess whether Sarah Stroud (2006) provides an account of epistemic partiality in friendship that is in conflict with evidentialism. In Stroud's 'Epistemic Partiality in Friendship', the term 'epistemic partiality' is used to refer to a bias in belief-formation and the content of beliefs. I agree with Stroud, argue that the two theories are incompatible, and locate the incompatibility: there is a conflict between the epistemic demands of friendship and the main principles of evidentialism.

I begin by outlining Stroud's argument for epistemic partiality and the account of evidentialism presented by Clifford (1877). I then address where the apparent conflict can be found. Following from this apparent identification of conflict, I then consider the three strongest objections to the claim that the epistemic demands of friendship are incompatible with evidentialism. I examine objections offered by Stroud (2006), Brown (2013) and Jollimore (2001), and then present responses to these objections. I conclude that they are unwarranted.

I conclude with a summary of the importance of this debate: if we accept that these two theories are incompatible, it arguably presents an important concern over whether evidentialism, alongside the other mainstream theories of epistemology, must be radically modified¹.

1 The Conflict

Before proceeding to the objections to the epistemic incompatibility of friendship and evidentialism, I will first give a brief overview of the main claims of the respective theories, then note where the conflict lies.

*Delivered at the BJUPS Summer Conference, 8 July 2017 at Wolfson College, University of Cambridge.

¹For one example, Stroud (2006, 518) argues that friendship is indispensable, a *necessity* in society. Due to its immense value, we should examine whether we must dismiss the mainstream theories of epistemology including evidentialism.

Stroud argues that friendship *qua* friendship places demands on our belief-formation processes, as well as the content of beliefs. The epistemic demands of friendship can be summarised as follows: first, friendship demands a bias on belief-formation. Stroud argues that an individual is more likely to question unflattering or negative evidence that is presented against a friend than he or she would question unflattering or negative evidence presented against a stranger; similarly, we are far more willing to undermine any negative beliefs presented about our friend; we easily attribute malicious and unsavoury motives to the stranger when presented with unflattering or negative evidence; lastly, we treat positive evidence of a friend as *prima facie* plausible, while less effort is directed in defence of a neutral party.

People also tend to construct alternative, positive interpretations or excuses of the unfavourable evidence. The explanation for this phenomenon is, as Stroud argues, that this takes a lot of mental effort, time and energy. Naturally we would not devote this much effort to non-friends.

The second epistemic demand of friendship is that friendship leads to different doxastic outcomes, in the sense that our actual beliefs are also biased. A partial belief-formation process leads to partial beliefs. Stroud claims that we are more likely to draw different conclusions from evidence if that evidence concerns friends, and are more likely to give alternative, negative explanations about a friend less weight than a positive one, even if it is less sound.

Given these preliminaries, I now turn to an explanation of evidentialism. For this essay, I will focus on the main form of evidentialism, proposed by William Clifford in 'Ethics of Belief'. Clifford's main epistemic principle is that it is always morally wrong to form a belief based on insufficient evidence. Clifford makes the normative claim that being epistemically blameworthy entails being morally blameworthy, and the epistemological claim, that beliefs should be proportionate to the available evidence. For Clifford, an epistemically safe (and therefore morally correct) belief, is formed only through sufficient evidence. Forming a belief based on insufficient evidence, or even sustaining an existing belief by intentionally ignoring doubts and avoiding evidence to the contrary, is not an epistemically 'safe' belief. Consequently, the demands of friendship are not in line with the epistemic standards of evidentialism: friendship requires an epistemic bias or epistemic irrationality, which conflicts with the requirement for beliefs only to be formed through sufficient reason.

2 The Appeal to Differential Information

Having now explained the two theories and where the conflict lies, we can now examine the objections that the two theories are in fact incompatible. One such objection, the 'appeal to differential information', is raised by Stroud (2006, p. 515). Stroud con-

siders that our doxastic practices in friendship are not actually biases: they are perfectly epistemically rational, for they are evidence-based. The apparent biases are in line with Clifford's main principle of evidentialism. If I form a favourable opinion about my friend, it is not because I am being epistemically impartial; I am using prior knowledge of my friend's personality, forming my judgement based on evidence of what I previously knew about their character. For example, if I heard that a close friend was supposedly rude and aggressive towards a charity collector, I would use my evidence of the character of my friend, and conclude the information must be wrong. I would, perhaps, consider what I know about her being a friendly person, and look back to times in which my friend gave money to charity. This reasoning process is not epistemically irrational, but epistemically justified in relation to evidentialism, for my belief has been formed upon sufficient, and crucially previously possessed evidence about my friend's character.

We have seen that Stroud has suggested a possible argument for why the demands of friendship are actually epistemically rational and compatible with evidentialism, and now should turn to look at her own response to the challenge.

Stroud's response to the appeal to differential information is based on the epistemic differences concerning friends and strangers. She argues that if we were given a body of favourable evidence about a stranger, and were then tested on how we responded to unfavourable news about them, we would not epistemically act in the same way we would with unfavourable news about a friend. She states that, for example, if I was a historian studying a figure I had a vast amount of favourable evidence on, and I was faced with the apparent information that the admirable historical figure had committed a crime, it is questionable if I would respond with as much hesitance, disbelief and overall partiality, as I would do when faced with the news that my friend committed a crime. In this example, the knowledge about the historical figure's actions is equal to the knowledge about the actions of my friend, and it is reasonable to assume that both bodies of information are equally reliable.

In line with Stroud, I very much doubt that I would go to the same lengths to scrutinise evidence, or attempt to think of alternative interpretations for the stranger as I would for a friend. This conclusion seems to be based on the fact that the individual is my *friend* and not a *stranger*. If this is so, then we can agree with Stroud that we are epistemically biased towards friends: our epistemic standards are not the same for friends as they are for strangers. From this, we can argue that the appeal to differential information no longer holds, though we could use prior knowledge to form beliefs about friends, we do not do so in the same way we would if we had prior knowledge about strangers: we treat our friends with a special partiality, however the acceptance of special partiality is in conflict with evidentialism.

Although I do agree with Stroud's response to the appeal to differential information,

the argument is convincing only if the intuition is accepted that an epistemic agent would not go to equal lengths to act epistemically partial towards a stranger as they would towards a friend. Whilst I agree with this intuition, I believe a further explanation must be given in defence of Stroud's response to the appeal to differential information, in order to strengthen the argument that the epistemic demands of friendship are not compatible with evidentialism. The next section of this paper offers such an explanation and develops Stroud's response further.

3 The 'Bias Blind Spot'

I will now turn to offer an explanation as to why we may actually share the intuition that we are biased towards friends, as Stroud claims, but fail to recognise this bias. My explanation is based on the psychological phenomena known as the 'Bias Blind Spot', a cognitive bias of recognising when *other* individuals are biased, but failing to recognise when you are biased (Pronin 2002). The bias blind spot is an unconscious process: people are unaware of how biased they are, and the influences it has on their decision-making and evaluation of evidence. Other work has suggested similar findings: people are generally unaware of how biased they are, due to the phenomena of *introspective illusion* (Ehrlinger 2005): people often inaccurately assess the origins of their mental states. Research on introspective illusion strengthens Stroud's response to differential information, for individuals claiming they do not act with bias towards their friends in comparison to strangers may fail to see that they act with bias. Consequently, the appeal to differential information is not successful in arguing that epistemic partiality in friendship and evidentialism are compatible. This is in conflict with the claims of evidentialism, which only allows for beliefs to be formed on sufficient reasons, different epistemic standards based on biases are not permissible.

4 'Epistemic Virtues'

Having considered the first objection to the incompatibility of friendship and evidentialism, I will now turn to address the second objection. Consider the argument put forward by Troy Jollimore (2001): he argues that Stroud is wrong to view the demands of friendship as epistemically irrational, for they can actually be viewed as *epistemic virtues*. One way to understand Jollimore's argument is in relation to an example of *epistemic attitudes*, which he presents in 'Love's Vision'. He argues that in an ideal world he would read every single philosophy paper with a high degree of attention and open-mindedness, doing his best to not let any presuppositions or personal views cloud his judgement. However, he claims this is impossible, for there are far too many philosophy papers to read, much less read them all with such an approach to detail.

Jollimore states that due to these constraints, he must be selective; therefore, he only reads his friend's papers with the necessary care. It is implausible to then claim that he would be epistemically irrational in approaching his friend's work with these epistemic attitudes, as in an ideal world he would approach every single paper in this way.

We can compare this example to Stroud's account of epistemic partiality. Jollimore claims the ideal way of approaching all philosophy papers with eagerness and a desire to consider the argument is just how friends approach evidence about another friend. Whilst Stroud claims that this behaviour is epistemically vicious, Jollimore believes it actually should be viewed as epistemically virtuous. He states that an ideal epistemic agent should approach every single epistemic situation with the open-mindedness and full attention reserved for friends.

If Jollimore is correct in his claims that epistemic partiality is epistemically virtuous, then evidentialism would *not* be incompatible with friendship. Such epistemic behaviours would arguably be in line with the claims that Clifford makes about belief-formation processes. The willingness that Jollimore describes would satisfy Clifford's demand. If Jollimore is correct that Stroud's epistemic findings concerning friendship are not epistemic vices but virtues, it follows that they are compatible with evidentialism².

However, the objection given by Jollimore does not succeed in successfully demonstrating how Stroud's theory is compatible with evidentialism, since Jollimore's interpretation of Stroud is questionable. Whilst Jollimore claims that Stroud exaggerates the epistemic behaviours of friendship, he underestimates Stroud's claims when she argues friends act with epistemic partiality. Jollimore's interpretation of the epistemic demands of friendship are too weak. Stroud goes further than to say that the display of epistemic behaviour towards our friends is to act with more open-mindedness and less judgement as Jollimore claims. For example, if we recall the epistemic behaviours of dismissing damning evidence, and putting in more effort to discredit factors that may portray friends in a negative way, we are aware that they are not examples of open-mindedness, but *biases*, which are not epistemically virtuous.

In summary, we have seen how Jollimore misunderstands Stroud's view of epistemic partiality by underestimating the extent to which friends act with bias and the strength of Stroud's claims. Therefore, it still stands that this behaviour is an epistemic threat,

²I am aware that virtue epistemology may be perceived as a radical departure from standard epistemology, and it is commonly perceived to be in conflict with evidentialism. However, arguments offered by alternative readings of virtue epistemology, such as those found in Helen Battaly's paper, 'Virtue and Vice, Moral and Epistemic' (2010), oppose this belief. A further examination into the alternative readings of virtue epistemology could support my claim that Jollimore's account of epistemic virtues can be perceived as compatible with the main claims of evidentialism.

and incompatible with the theory of evidentialism.

5 Scientific Practice

I will now address an objection raised by Curtis Brown: he argues that the epistemic behaviours Stroud describes are not as epistemically irrational or contradict evidentialism. Brown does not state that the demands are epistemic virtues; instead, he argues epistemic partiality is in line with attitudes towards the formation of scientific theories. The *origins* of a scientific theory do not matter in assessing the theory, only whether it is confirmed or disconfirmed. This norm of scientific practice permits forming a belief based on biases, even the desire to believe the best about a friend³.

As this behaviour is epistemically rational in scientific practice, it is equally epistemically rational when forming beliefs about friends.

Assuming Brown's conclusions, Clifford's evidentialism can be understood to allow for hypotheses to be formed upon insufficient evidence. If Brown is correct in rationally explaining Stroud's epistemic behaviours in friendship, Stroud's epistemic partiality in friendship is at least partly compatible with evidentialism.

However, the key problem with Brown's argument is that even if evidentialism did permit provisional acceptance or entertaining hypotheses to be formed through partiality, it does not permit *beliefs* to be formed in this way. Friendship involves forming biased beliefs, not provisional acceptance or entertaining hypotheses. Brown himself argues that we would not be critical of a scientist if they favoured a hypothesis, as long as they appropriately formed a belief based on the evidence. However, Stroud has argued that unlike scientist, a friend does not view evidence or findings with impartiality, and they are more likely to only credit positive 'data' towards their findings. Brown therefore fails to successfully argue that Stroud's epistemic findings of partiality in friendship can be deemed as in accordance with scientific practice. At most he explains how friends are allowed initial partiality with what they inquire about friends, just as a scientist is allowed to form favourable hypotheses; however, his argument does not extend to the formation of beliefs. Even if evidentialism would allow for insufficient hypothesis-formation, it does not allow for insufficient belief-formation. Therefore, despite Brown's objection, the two theories are not compatible, as a conflict is still present between evidentialism and Stroud's account of epistemic partiality.

³An argument offered by Patrick Maher (2006) is similar to Brown's claims. Maher argues that scientists *provisionally accept* or *entertain* hypotheses for testing, and form beliefs about them once they have been corroborated by empirical evidence.

6 Conclusion

In conclusion, this paper has considered three objections against the compatibility of Stroud's epistemic partiality and Clifford's account of evidentialism. I have argued the two theories are incompatible: a real conflict is present between the epistemic demands of friendship and the main principles of evidentialism. The next step would be to determine whether friendship or evidentialism should be accepted.

References

- [1] Battaly, H. (Ed.) (2010). *Virtue and Vice, Moral and Epistemic*. Oxford: Wiley Blackwell.
- [2] Brown, C. (2013). 'Friendships: Epistemically Dangerous Liaisons?' Caluori, D. (ed.). *Thinking about Friendship, Historical and Contemporary Philosophical Perspectives*. Palgrave Macmillan: 99–117.
- [3] Campbell, J.D. (1986). 'Similarity and uniqueness: The effects of attribute type, relevance, and individual differences in self-esteem and depression'. *American Psychological Association*.
- [4] Chignell, A. (2010). 'The Ethics of Belief' *Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/ethics-belief/>.
- [5] Clifford, W.K. (1877) [1999]. 'The ethics of belief', in T. Madigan (ed.), *The ethics of belief and other essays*. Amherst, MA: Prometheus: 70–96.
- [6] Ehrlinger, J, Gilovich, T & Ross, L. (2005). 'Peering into the bias blind spot: People's assessment of bias in themselves and others'. *Personality and Social Psychology Bulletin*, 31(5): 680–692.
- [7] Jollimore, T. (2001). *Love's Vision*. Princeton: Princeton University Press.
- [8] Keller, S. (2010). 'Friendship and Belief'. *Philosophical Papers*, 33(3): 329–351.
- [9] Maher, P. (2006). Confirmation Theory. In: Borchert, D, ed. 2006. *The Encyclopedia of Philosophy* (2nd edition). Macmillan Reference, USA.
- [10] Pronin, E. (2002). 'The Bias Blind Spot: Perceptions of Bias in Self Versus Others' *SAGE Journals*. <http://psp.sagepub.com/content/28/3/369.short>
- [11] Stroud, S. (2006). 'Epistemic partiality in friendship'. *Ethics*. 116(3): 498–524.

A Marxist Review of Justice

Giulia Carlini

University College London

Abstract

I investigate whether Karl Marx's contrastive critique of capitalism has a normative dimension; I conclude Marx's critique does not entail capitalism is unjust. Though there are no explicit normative claims in his writings, Marx's description of capitalism (including his choice of words) suggests Marx nevertheless believes capitalism is unjust. I address this inference by first considering the claim, 'capitalism is unjust because it hinges on unequal transactions between the capitalist and the worker' in a Marxian framework (§1). I show: (a) the transaction is equal; (b) even if it were unequal, capitalism is not unjust based on Marx's critique because Marx viewed morality (an ideology), as being constructed by the ruling class in their best interests. It follows that Marx holds capitalism to be just *because* it is in the interest of the capitalist class for capitalism to hinge on an unequal transaction between themselves and the proletariat class. In (§2) I present Ziyad Husami's case for a double determination of morality which accounts for a proletarian theory of justice (denouncing capitalism as unjust) and argue that it is a plausible, unfounded solution. Last, in (§3), I attempt to account for Marx's recurring use of seemingly edifying words. I conclude that within a Marxist framework, capitalism is just.

1 Equal and unequal transactions in a Marxian framework

Marx speaks of the circumstances under which the working class produces its livelihood, i.e. the exchange of labour power for wages. In capitalism, manual labour power is traded like any other commodity in alignment with Ricardo's labour theory of value: 'The value of a commodity, or the quantity of any other commodity for which it will exchange, depends on the relative quantity of labour [time] which is necessary for its production' (Ricardo 1821, §1.1). Marx changes labour to the *socially* necessary labour time, i.e. the average labour of average productivity. According to Marx, wages will correspond to the quantity of labour that is socially necessary to produce the worker: his livelihood and the cost that the capitalist would incur to replace him were he to die or quit (Marx 1969, §1.6.11). However, the worker is then not trading his labour time, but rather his labour-power: his *ability* to do work. Thus, per Marx's labour theory of value, the worker is paid only the 'necessary labour' to reproduce his wage (ibid., §1.7.3) As the efforts of the worker will yield more than what he is compensated for, the capitalist is able to profit by keeping the surplus value: the difference between the

value of labour power and the cost of socially necessary labour time (in wages) (ibid., §1.8.5). Given this reading, we might be inclined to think (A):

- (P₁) This transaction is unequal.
- (P₂) This transaction is unequal.
- (C) Therefore, this transaction is unjust.

Taking (A) one step further leads to the conclusion, ‘capitalism is unjust because it is exploitative’. Marx’s choice of words seems to affirm this inference: for example, when he speaks of the ‘*exploitation* of some by others’ (Marx 2008, 18) and of the ‘*robbery*’ of the worker (Marx 1887, §1.15.4.9) he deliberately chooses words that can very easily imply concepts of unfairness and injustice (B) – this cannot be coincidental.

I now show that, although tempting, A does not imply Marx viewed capitalism as unjust. Namely, in a Marxist framework P₁ and P₂ are false; consequently, C is unsound. P₁ is false since Marx argues the transaction is equal: every commodity has an intrinsic value that corresponds to the quantity of socially necessary labour-power needed to produce it; labour-power is a commodity; therefore, the value of labour-power is the worker’s livelihood. The capitalist, in consuming this commodity, has the right to extract all the value he wishes, resulting in the worker labouring more than what is necessary to reproduce his wage. It follows that Marx, in considering labour-power a commodity, does not hold P₁ to be true. Even if we assume Marx considers this transaction as being unequal (as the capitalist appropriates the surplus value), where would that leave his moral evaluation of capitalism? If, for example, Marx’s conception of justice were distributive (P₂) then C would follow. We (including Marx) might consider the transaction as being unjust on, for instance, utilitarian grounds: this transaction does not maximise utility (Driver 2009) so it is unjust. Yet, Marx’s historical materialism suggests the opposite: capitalism is just *in virtue* of its exploitative nature.

Marx rejects all political, juridical and spiritual conceptions of history – in their stead he offers a conception that is both materialistic and economic: he reduces history to the development of human productive activity. To understand history is to fathom the material conditions in which mankind produced the means of its subsistence. Its genesis, Marx contends, can be traced back to the moment in which technology allowed for a surplus. This made it possible for some to live off others. The motor of history is then a dialectical antagonism: it sees the ruling class dictating the circumstances of labouring and the working class living in it (Marx 1859): ‘The history of all hitherto existing society is the history of class struggles’ (Marx 2008, 3). The economic structure underpinning history, Marx continues, is comprised of:

- (1) ‘Productive forces’: the elements necessary for the productive activity (including labour-power, the means of production and the technical

and scientific knowledge required for the organisation and advancement of production).

(2) ‘Relations of production’: the social relations that men have to involuntarily participate in to secure their livelihood; these determine the ownership and use of the productive forces, as well as the allocation of their products.

Each society proceeds to a new stage through class struggle and revolution, not gentle evolution. That is because any economic system necessarily stimulates more growth than it can contain: ‘like the sorcerer who is no longer able to control the powers of the nether world whom he has called up by his spells’ (ibid., 8). As the productive forces grow more social, the static relations of production begin to fetter their growth; the working class then revolts against the ruling class to eventually replace it and in doing so it determines a whole new economic system (ibid., 9). Now, the new economic structure will go on to determine and condition *every* aspect of society: the institutions that form a state (juridical, political, cultural, religious, etc.) as well as, indirectly, any ideas that are made prominent in a society with these very institutions (Marx 1859, 6–7). This suggests that the concept of justice is too determined and conditioned by the relations of production (economic system), so will be in alignment with whichever best suits the interests of the ruling class (maximisation of profits). As the exploitation of the worker is necessary for the functioning of the capitalist system and the ruling class develops the conception of justice, not only is exploitation *per se* just, but any attempt to resist it would qualify as an injustice (Wood 1972, 268).

Capitalism is just *because* it is exploitative. Thus, claiming that the capitalist exploitation of the worker is a form of injustice because for instance it (a) does not distribute fairly the product of labour or (b) does not maximise utility (in other words does not abide to a distributive or utilitarian principle of justice), is misguided: the standard by which is it possible to be morally critical of society is dictated by the ruling class; in other words, there is no such thing as an independent critical perspective. Hence, following Marx’s findings, any attempt to adopt an external moral standpoint (that could abide to a distributive, utilitarian or any other conception of justice) will necessarily result in incorrect assumptions. This should provide sufficient ground to claim that Marx would reject (A) altogether, and instead claim that capitalism is just *because* it is exploitative.

2 Husami’s case for double determination of morality

Husami disagrees with Wood’s interpretation of Marx’s morality and argues it is possible to assess a society’s level of justice by adopting a moral standard that is not the ruling class’s; namely, by adopting one developed by the working class in their own

interest (Husami 1978, 33). The working class, in applying its own concept of justice to the capitalist system of exploitation, arrives at a different evaluation: capitalism is unjust *because* it is exploitative. Then, following these findings, a society is split by two moralities (and corollary theories of justice) that respect the interests and needs of the two opposing classes: following Husami's reasoning, while the bourgeois morality will maintain its hegemony, it does not preclude the existence of a proletarian morality. Now how does this finding apply to our investigation of Marx's moral evaluation of capitalism? If we were to accept Husami's account, it follows that either (a) Marx considers himself a member of the bourgeois and would adopt the bourgeois theory of justice which considers capitalism a justice or (b) Marx considers himself a proletariat and would adopt the proletarian theory of justice which considers capitalism an injustice. *The Communist Manifesto* provides a strong case for (b):

The Communists, therefore, are on the one hand, practically, the most resolute section of the working-class parties of every country, that section which pushes forward all others; on the other hand, theoretically, they have over the great mass of the proletariat the advantage of clearly understanding the line of march, the conditions, and the ultimate general results of the proletarian movement (Marx 2008, 17).

This passage suggests that Communism is an expression of the proletarian movement, thus Communist theorists, like Marx, identify with the working class. This challenges the interpretation of his moral theory as a form of 'vulgarly determinism' (which is an ideology *absolutely* determined by class background) and thus makes some room for self-determinism.

However plausible Husami's claim is that there's a double determination of morality by which it is possible to evaluate a society (in this case as simultaneously just and unjust), Marx provides no textual evidence to support this. If Marx did adopt a proletarian moral standpoint (and consequently condemned capitalism) then he goes to great lengths to hide it. We could speculate that he does not want to go out of his way to say that capitalism is unjust because, as Tucker suggests:

The distributive orientation ultimately pointed the way to abandonment of the revolutionary goal (Tucker 1969, 48).

In other words, assuming there exists a proletarian theory of justice (that could be distributive, or perhaps even better, utilitarian) openly appealing to it would have persuaded workers to compromise and reform rather than revolt, which Wood argues would have compromised the transformation to socialism and then communism (Wood 1972, 274). Wood argues Marx refrained from appealing to a theory of justice *not* because it sends the wrong message (reformation over revolution), but because it fundamentally presupposes a theory of revolution he believes is disproved. Hence we can appreciate the significance of Husami's argument: it would (a) prove that Marx

did consider capitalism to be unjust and (b) explain why Marx refrained from openly stating it. However, Husami fails to account for Marx's omission of a second moral determination. This leaves the Husami's case for a double determination entirely unfounded and so it falls short of satisfactorily ascribing to Marx the view that capitalism is unjust.

3 Marx's edifying words

I have argued Marx does not consider capitalism to be unjust per A. Now, I consider B: Marx's choice of highly suggestive words (e.g. the aforementioned exploitation and robbery, etc.). I argue these are part of his rhetorical style and expression, which is aimed at prompting the proletariat to revolt and so do not imply a moral dimension to his materialist account. In the *Manifesto*, we can pinpoint many occasions in which Marx makes uses highly evocative expressions, the most blatant being the following: 'The proletarians have nothing to lose but their chains. They have a world to win. Working men of all countries, unite!' (Marx 2008, 39). Words as such are intended to pick up momentum and inspire change (especially considering the revolutionary expectations of the 1840s) (ibid, 12) and therefore do not necessarily carry moral value.

This essay considered the claim that Marx considers capitalism to be unjust although his works lack in explicit normative claims. It argued Marx does not consider capitalism to be unjust, but rather the opposite. It did so by rejecting the claim that 'capitalism is unjust because it hinges on unequal transactions between the capitalist and the worker'. I showed this claim is incompatible with Marx's account because (a) Marx believes the transaction is equal and (b) even if it were unequal, capitalism would be just because Marx viewed morality as a dictate of the ruling class (and so best accommodates their interests). First, I showed (I) it follows Marx must consider capitalism to be just *in virtue* of its exploitative nature; then, I (II) presented the case for the existence of a double determination of morality, as suggested by Husami, and reflected on the significance of this claim. However, I rejected Husami's case for he neither provides evidence to support it, nor puts forward a satisfactory account for this lack of evidence. Last, I (III) accounted for Marx's choice of highly suggestive words (*robbery, exploitation, theft*) without appealing to morality and thus we are left with a claim difficult to resist: Marx does not view capitalism as unjust.

References

- [1] Driver, J. (2014). ‘The History Of Utilitarianism’. In *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), Winter 2014. Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/win2014/entries/utilitarianism-history>.
- [2] Husami, Z.I. (2008). ‘Marx On Distributive Justice’. *Philosophy & Public Affairs* 8(1): 27–64.
- [3] Marx, K. (2008). *The Communist Manifesto*. Oxford: Oxford University Press.
- [4] Marx, K., S. Moore, and E. Aveling. (1887). *Capital*, Vol. I. 1st ed. Moscow: Progress Publishers.
- [5] Marx, K. (1859). *A Contribution To The Critique Of Political Economy: Preface*. 1st ed.
- [6] Marx, K. (1969). *Value, Price And Profit*. 1st ed. New York: International Co.
- [7] Ricardo, D. (1821). *On The Principles Of Political Economy And Taxation*. 3rd ed. London: John Murray.
- [8] Tucker, R. C. (1969). *The Marxian Revolutionary Idea*. New York: Norton.
- [9] Wood, A. W. (1972). ‘The Marxian Critique Of Justice’. *Philosophy & Public Affairs* 1(3): 244–282.

Would a Life Without Death Be Intolerable?

Joshua Goh

University College London

Introduction

In this essay, I argue a life without death, characterised in a particular manner, would likely be intolerable.¹ This essay has five parts. The first section characterises ‘a life without death’ (hereafter referred to as ‘immortality’). The second section formally presents my argument: in brief, immortality would likely eventually become intolerable, *insofar as it would eventually become meaningless*. The third section articulates the strongest objections to the most controversial premise of my argument. The fourth section defends this premise against an objection. Lastly, I summarise the argument.

1 Characterising Immortality

Borrowing from John Martin Fischer and Benjamin Mitchell-Yellin (2004, 363–365), the kind of immortality I wish to consider is as follows:

At a somewhat felicitous age (say 36), Albert’s body is ‘biologically frozen’. While still vulnerable to various contingencies, it is no longer subject to biological deterioration, or the contraction of fatal disease and illness. Also, Albert’s psychological continuity is preserved for as long as he lives.²

Two preliminary notes on the kind of immortality are necessary. First: I am considering ‘medical’, and not ‘true’, immortality. As Fischer and Mitchell-Yellin note (2014, 363), Albert can be either ‘medically’ immortal or ‘truly’ immortal. The former is characterised as above whereas the latter means that Albert is invulnerable to death by *any* cause (ibid., 363–365). That is to say, Albert would not die, even if he fell off a skyscraper. I consider medical, and not true, immortality, for two reasons:

¹Bernard Williams argues that a life without death, ‘where conceivable at all’, must *necessarily* be intolerable (1973, 82). My claim is weaker.

²The immortality of Elina Makropulos (EM) seems to be of this kind.

First, medical immortality is much easier to conceive of. Thus, it is less vulnerable to the charge that, because immortality is so inconceivable, debates about its value become conceptually incoherent. Mikel Burley (2009) levels this charge; Timothy Chapell responds (2010, 105–106).

Secondly, debates about medical immortality yield more insight into practical matters. For example, Donald Bruckner (2012, 625–626) notes that this debate contributes to the discussion about the value of life-extending therapy. In contrast, debates about true immortality yields hardly any insight into practical matters.

A second preliminary note: I am considering an immortality which involves *psychological continuity*. Thus, I have little to say about, for example, an ‘indefinite [...] series of psychologically disjoint lives’ (Williams 1973, 92).³ Nevertheless, I grant that Williams is right to question how sensible it would be for Albert to desire such an (psychologically disjointed) immortality (ibid., 91). However, this is tangential to whether or not said immortality must be intolerable, *should Albert choose it*.

In any case, all references to ‘immortality’ now refer to the above-stated, psychologically continuous, medical, immortality (PCMI).

2 The Formal Argument

I will first explain the notion of a ‘categorical’ desire before presenting my formal argument.

A categorical desire differs from a conditional desire (ibid., 85–86). A conditional desire is a desire that is conditional on continuing one’s own life. But it isn’t *reason for* continuing one’s own life. Samuel Schefferler (2013, 88) gives the following example to illustrate conditional desires: as Albert continues life, he desires his tooth cavities filled. But this desire doesn’t give Albert reason to continue living. On his deathbed, Albert wouldn’t think: ‘I wish I had more time to get my cavities filled’.

In contrast, a categorical desire is not conditional like so. It is a desire for something (call this something a ‘categorically desirable’ thing) which makes life not just bearable, but worth continuing. Categorical desires, then, ‘[propel us] forward into longer life’ (Williams 1973, 91). That is, they give us reason to resist death, which would frustrate these desires. On his deathbed, Albert might think: ‘I wish I had more time to be with my granddaughter’.

³Cf. A. W. Moore’s ‘The Repeating Life’ (2006, 316–318).

As David Beglin (2016, 14) notes, categorical desires drive our various, meaningful, life commitments. As an illustration, Albert categorically desires meaningful personal relationships, intellectual, physical, and artistic accomplishment, and social justice.⁴ Consequently, he commits his life towards cultivating, or appreciating, these things.

I now formally present the argument:

(P1) If we lose our categorical desires, then we will lose reason to live.

(P2) If we are immortal, it is likely that we will eventually lose our categorical desires.

(C1) If we are immortal, it is likely that we will eventually lose reason to live. (P1 + P2)

(P3) If we have no reason to live, then life will eventually become intolerable, insofar as it becomes meaningless.

(C2) If we are immortal, it is likely that life will eventually become intolerable, insofar as it becomes meaningless. (C1 + P3)

Subsequently, I focus on articulating, and defending, what I think is my most controversial premise: (P2). Nevertheless, I now talk a little about the other steps of my argument.

First, consider the inferences from (P1 + P2) to (C1), and from (C1 + P3) to (C2). These are both valid, by hypothetical syllogism.

Secondly, consider (P1). Perhaps one might have reason to live, even in the absence of categorical desires. For example: Albert deems that life has nothing left to offer him. However, he perceives God telling him to wait for the second coming of Jesus Christ. Albert has reason to live; he is petrified of that divine retribution he might face, should he disobey.

Thus, we might resist death, but for reasons divorced from categorical desires. Nevertheless, I think it a reasonable assumption that, for most, categorical desires, if not *exclusively* what give us reason to live, constitute enough of said reason, such that, in their absence, only a negligible amount of said reason remains.

Lastly, consider (P3). Perhaps a meaningless life doesn't need to be completely intolerable. Albert, I suppose, might find a life of meaningless sensual pleasure quite comfortable.

Now, a meaningless *mortal* life already seems, to me, quite intolerable. But some may disagree. To them, I reply: the meaninglessness that must ultimately be considered

⁴I admit – it is debatable how, or even if, Williams would classify (truly) impersonal, 'non-I' (1973b, 260-261), desires as categorical.

here is an *eternal* meaninglessness. In any case, (P₃) is (unabashedly) a *pro tanto* premise; its strength (and, by extension, that of (C₂)) depends on how intolerable said eternal meaninglessness is judged to be.

However, suppose that one is not convinced by (P₁) and/or (P₃). Nevertheless, in establishing (P₂), I will have shown, at the very least, that: if immortal, we would lose *a good deal* of our reason to live. Whether *all* of our reason to live is lost, and whether or not such a meaningless state is *completely* intolerable, is open to further debate.

One last note: I am not arguing that, given the choice, one should not choose to be immortal. My thesis can ultimately be reconciled with a choice to become immortal. After all, given our characterisation of immortality, one might choose to become immortal, but then (somehow) opt out of said immortality, *once it becomes intolerable*. Williams (1973, 100) may endorse such a plan.

3 Articulating reasons for (P₂)

(P₂) If we are immortal, it is likely that we will eventually lose our categorical desires.

I have two reasons for (P₂).

- (1) We have little reason to think that our categorical desires could be eternally sustained.
- (2) We have good reason to think that, over time, our categorical desires would diminish.

I articulate (1) and (2) in turn.

(1) We have little reason to think that our categorical desires could be eternally sustained

Now, could we clearly conceive of an eternally categorically desirable thing (call this thing a (1)-defeater), then we'd have some reason to think that our categorical desires could be eternally sustained.

What might be a defeater for (1)?

Contrary to Williams, it need not be 'something that makes boredom unthinkable ... something that could be guaranteed to be at every moment utterly absorbing' (ibid., 95).

Fischer is correct here: in those mortal lives worth living, we don't require that categorically desirable things engage one *at every instant*. Why think that this requirement

changes in an immortal life? It isn't as though, when immortal, our need to be engaged increases so dramatically, such that even a millisecond of boredom would make life not worth living (Fischer 1994, 261–262).

Let us lower the bar – a (1)-defeater need only be something that Albert could experience infinitely many times, but still find worth living for. Albert can occasionally tire of this (1)-defeater, but his desire for it must forever rekindle. Can we clearly conceive of such a thing?

Fischer's 'repeatable pleasures' (hereafter 'RPs') offer themselves as candidates. These are those pleasures which one finds 'highly fulfilling and completely satisfying at the moment', and which one wishes 'to have more [of] ... at some point in the future' (ibid., 263).

As Fischer and Mitchell-Yellin (2014, 356) note, RPs are many, and diverse. There are 'self-focused' RPs, like the 'the pleasures of eating delicious food, listening to beautiful music, appreciating great art, and having sex'. And there are also 'other-focused' RPs, like the pleasures of '[finding] cures for terrible diseases and health problems ... , understanding the underlying physics of the universe, [and] writing ... great novels or beautiful poetry' (ibid., 357–358).

To Fischer and Mitchell-Yellin, at least one RP might conceivably be a (1)-defeater. After all, there are just so many RPs which they cannot imagine ever tiring of.

I do not, however, think that any RP is clearly conceivable as a (1)-defeater.

For one thing, I doubt there is that strong an intuition about the eternal repeatability of any RP. Now, Fischer says that he personally has 'no tendency to think that [many RPs] would become less compelling, unless pursued in a single-minded or compulsive fashion' (Fischer 1994, 266).

However, while some philosophers, including Fischer, may have this intuition, it is not clear that most do. For every philosopher that shares Fischer's intuition, you will find a philosopher like Shelly Kagan, who loves philosophy, math, and cinema. However, to Kagan, the prospect of repeating any of these pleasures (even in what Fischer would deem an appropriate fashion), 'not just for years, or decades, or even centuries, but for all eternity ... [is] a nightmare' (Kagan 2012, 240).

Let me emphasise that to clearly conceive of an RP as a (1)-defeater, it isn't enough to show it somewhat, or even very, repeatable. One must show it eternally repeatable.

Alas, even if most did match Fischer's appetite for a particular RP, this hardly shows that said RP is as such. After all, that we have *so far* not tired of an RP, is hardly quantitative evidence for the claim that we will never, *ever*, tire of the said RP. As compared to eternity, the time period of a finite life is simply too short to be the basis upon which one reaches a conclusion about our infinite capacity for an RP.

Thus, I doubt that any RP is clearly conceivable as a (1)-defeater. And, in the absence of any (1)-defeater, (1) plausibly stands. Let's now look at (2).

(2) We have good reason to think that, over time, our categorical desires would diminish

(2) is motivated by this following thought: many of the things that we find categorically desirable, remain thus so, only against a backdrop of finitude.

Now, on one level, categorically desirable things are thus so because of their experiential value. When we cultivate, or appreciate, them, it feels good.

However, on another level, many categorically desirable things are thus so because we impart onto them a certain *symbolic* value (Beglin 2016, 13). That is, we live for these things, because we take them to *mean something*, or to be *personally significant*.

But what exactly do their 'meaning something', and being 'personally significant', involve? I suggest, along with Beglin, the following: we have a scarce amount of time. Thus, when Albert decides to invest said time in, say, athletic accomplishment, athletic accomplishment takes on a *relative* significance, 'relative to the other values and projects [Albert] won't be making central to [his] life. In a sense, [Albert's] life comes to *stand for* [athletic accomplishment]' (ibid., 12, emphasis added).

However, in an immortal life (devoid of temporal scarcity), categorically desirable things could no longer possess said relative significance. And neither would our pursuing them incur any opportunity cost.

As Martha Nussbaum notes, gone would be that 'intensity and dedication' with which mortal Albert had treated his time at the gym (Nussbaum 1994, 229). After all, Albert now has all of eternity both to set the world sprint record, and to do anything else he pleased⁵.

Let's now turn to the examples of deep personal relationships. Aside from their being experientially satisfying, much of their significance, and the urgency with which we treat them, stems from the parties involved each having considerably invested their most precious, limited, resource (time) into one another.

But say we were immortal. Then there would no longer exist the sense in which a relationship is significant because one has, *at the certain expense of other (potential)*

⁵Scheffler, in fact, thinks that Albert might lose *all* concept of value (2013, 98–100).

relationships, chosen *this* relationship to be central to one's life⁶.

In a somewhat different vein, I want to add that: many of those categorically desirable things which we *appreciate*, say, physical and artistic accomplishment, require for their appreciation a backdrop of finitude. For example, how significant would an athlete's winning a tournament be, if she had all the time in the world to compete? Not very significant at all.

Furthermore, all that film, art, and music, which have as their content either the tragedy or celebration of mortality, transitoriness, aging, and/or death, would lose much of their emotional weight. Think of, for example, Michael Haneke's *Amour* (2012), Eagle Eye Cherry's *Save Tonight* (1997), Taylor Swift's *22* (2013), and Caravaggio's *Young Sick Bacchus* (1593-1594).

Thus, immortality robs many categorically desirable things of their symbolic value, and many other categorically desirable things of their appreciative potential. And, because said symbolic value, and appreciative potential, are a large part of what makes these things categorically desirable, (2) plausibly stands.

Combining (2) with (1), then, (P₂) seems quite likely, if not necessarily, to be true. We have, in the end, more reason than not to believe that: 'if immortal, we will eventually lose our categorical desires.'

4 Defending (P₂)

Objection: Even if no RP is clearly conceivable as a (1)-defeater (as *eternally* categorically desirable), perhaps our categorical desires could still be eternally sustained.

Imagine that Albert gets tired of RP *x*. However, he discovers RP *y*, and he deems it worth living for. And, when Albert tires of RP *y*, in comes RP *z*; ad infinitum.

Thus, one's categorical desires might be eternally sustained by way of an indefinite replacement of RPs that are, at least, provisionally categorically desirable (call these PCD-RPs).

This replacement, of course, could only be indefinite, if our stock of PCD-RPs is infinitely vast. And perhaps it might indeed be 'indefinitely and incalculably large' (Chappell 2007, 37), and not 'finite and fixed' (Fischer and Mitchell-Yellin 2014, 360).

For one thing, there seems an infinitely vast stock of self-focused RPs from which we might draw PCD-RPs. Bruckner, for example, argues 'human ingenuity and invention'

⁶Beglin (2016, 10–11) and May (2009, 63) express similar ideas.

has, and will continue to, generate new, engaging, and previously unimaginable, activities (Bruckner 2012, 632–634; Wisniewski 2005, 33–35). A few centuries ago, who might have imagined Albert watching Netflix, and playing Minecraft? We should, moreover, expect this inventive trend to continue indefinitely.

For another thing, this also seems to be the case for other-focused RPs. For example, Fischer and Mitchell-Yellin (2014, 358) doubt ‘there will *ever* be a time when we will have banished all serious health challenges that would generate projects that could propel people into the future’.

Response: One first worry: we each possess a somewhat fixed character (a somewhat stable set of dispositions, values, interests, and goals).

Thus, there might exist an infinitely vast stock of RPs ‘out there’. However, we can only draw PCD-RPs from that (limited) set of RPs which matches our character. A recluse, for example, could never enjoy five days of Glastonbury, and a socialite could never enjoy meditative solitude. To think that they might seems, as Williams writes, an attempt to ‘combine the uncombinable’ (Williams 1973, 94).

But perhaps drastic character change is possible, especially over as long a time as eternity⁷. Thus, one might (eventually) draw from the entire set of RPs ‘out there’.

However, even so, I still don’t think it plausible that an immortal’s stock of PCD-RPs can be infinitely vast. This is for two reasons.

First, I simply express the sentiment: many of Bruckner’s examples do not seem engaging enough to qualify as PCD-RPs. Certainly, ‘[riding] a bicycle’, and ‘a game of *Trivial Pursuit* with friends’, may ‘keep us from getting bored’ (Bruckner 2012, 634)). However, and although this might reflect personal prejudice, I’m not sure if these activities are ones we would *live for*.

Second, we have, in any case, more to minus from an immortal’s stock of PCD-RPs. As per section 3, we should minus from this stock both those RPs that require for their appreciation a backdrop of finitude, and those RPs that we live for in virtue of their symbolic value.

However, of course, some self-focused RPs are engaging enough to live for, and do not require for their appreciation a backdrop of finitude (e.g. listening to Schubert).

⁷Williams (1973) argues: it is unclear how attracted one could be to an immortality involving such character change. After all, how much might a nun *presently* care for a future life of orgiastic pleasure? However, even if, here, Williams is correct. (Cf. Bruckner (2012, 637–639) and Fischer (1994, 267–268) for why Williams may not be correct.) This gives us little reason to think, *should said nun choose said immortality*, and *should said character change occur*, *said immortality must be intolerable*.

Moreover, I doubt not that there will, in the future, emerge music at least as engaging as *Ave Maria*).

And, indeed, some other-focused RPs should remain significant in some sense, even if they lose that relative significance abovementioned (e.g. achieving medical cures, and promoting social justice).

However, still I doubt that these RPs are enough to constitute an *infinitely* vast stock of PCD-RPs.

For one thing, I think that even those self-focused RPs, which remain immensely engaging in an immortal life, will eventually not be worth living for.

This is because we aren't rats. We are creatures who 'reflect on our first-order or base-level experiences' (Kagan 2012, 242). And, like Kagan, I think that said reflection will eventually force the question: 'is this really all there is to life, just simple pleasure like this?' (ibid.)

Thus, there comes a point where an immortal finds it difficult to *relate meaningfully* to self-focused RPs, pleasurable as they might be. Simply put – pleasure is no longer enough. At this point, RPs stop motivating continued existence, and one's stock of self-focused PCD-RPs is empty⁸.

But what of those other-focused RPs which conceivably retain some sort of intrinsic significance, even if they lose relative significance? Plausibly, one might pursue social justice indefinitely, without ever thinking, 'what is the point of this?' Said intrinsic significance is also reflected, as Beglin (2012, 15) writes, by how 'we can imagine marveling at someone's eternal dedication to promoting justice'.

I concede – for one disposed to such other-focused RPs, it might just be that her categorical desires could be eternally sustained. (See (Beglin 2012, 15–17).) Alas, I am not sure that many actually want to spend an eternity promoting social justice or achieving medical cures. In reference to that latest Beglin quote, I think that a large part of our marveling derives from how we cannot imagine ourselves making a similar dedication!

In summation, we should doubt that an immortal's stock of PCD-RPs can be infinitely vast. And, thus, we should doubt that one's categorical desires might be eternally sustained by way of an indefinite replacement of PCD-RPs.

⁸Here, one might say that we succumb to Williams's particular conception of 'boredom' (1973, 95).

5 Conclusion

I have argued (1) we have little reason to think that our categorical desires could be eternally sustained, and that (2) we have good reason to think that, over eternity, our categorical desires would diminish. Furthermore, (3) we should doubt that one's categorical desires might be eternally sustained by way of an indefinite replacement of PCD-RPs . Thus, (P2) if immortal, it is likely that we will eventually lose our categorical desires.

Granted that (P1) categorical desires are just what give us reason to live, and that (P3) a meaningless life is regarded as intolerable, this suggests that: a life without death, characterised as PCMI , would, likely (and eventually), become intolerable, insofar as it would eventually become meaningless.

References

- [1] Beglin, D. (forthcoming). 'Should I Choose to Never Die? Williams, Boredom, and the Significance of Mortality', *Philosophical Studies*: 1–20. Available from: <http://link.springer.com/article/10.1007%2F978-0-16-0785-1> [Last Accessed 7th January 2017].
- [2] Bruckner, D. W. (2012). 'Against the Tedium of Immortality', *International Journal of Philosophical Studies*, 20(5): 623–644.
- [3] Burley, M. (2009). 'Immortality and Meaning: Reflections on the Makropulos Debate', *Philosophy*, 84(4): 529–547.
- [4] Chappell, T. (2007). 'Infinity Goes Up on Trial: Must Immortality Be Meaningless?', *European Journal of Philosophy*, 17(1): 30–44.
- [5] Chappell, T. (2010). "'A Logos that Increases Itself': Response to Burley", *Philosophy*, 85(1): 105–108.
- [6] Fischer, J. M. (1994). 'Why Immortality is Not So Bad', *International Journal of Philosophical Studies*, 2(2): 257–270.
- [7] Fischer, J. M. and B. Mitchell-Yellin (2014). 'Immortality and Boredom', *Journal of Ethics*, 18(4): 353–372.
- [8] Kagan, S. (2012). *Death*. Yale: Yale University Press.
- [9] May, T. (2009). *Death (The Art of Living)*. Stocksfield: Acumen Publishing Limited.
- [10] Moore, A. W. (2006). 'Williams, Nietzsche, and the Meaninglessness of Immortality', *Mind*, 115(458): 311–330.
- [11] Nussbaum, M. (1994). *The Therapy of Desire*. Princeton: Princeton University Press.
- [12] Scheffler, S. (2013). *Death and the Afterlife*. Oxford: Oxford University Press.
- [13] Wisniewski, J. (2005). 'Is the Immortal Life Worth Living?', *International Journal for Philosophy of Religion*, 58(1): 27–36.
- [14] Williams, B. (1973). 'The Makropulos Case: Reflections on the Tedium of Immortality', in *Problems of the Self: Philosophical Papers*. Cambridge: Cambridge University Press.
- [15] Williams, B. (1973b). 'Egoism and Altruism', in *Problems of the Self: Philosophical Papers*. Cambridge: Cambridge University Press.

Solutions to the Old Evidence Problem in Bayesian Confirmation Theory*

Matthew Coates

Queen Mary University of London

Introduction

At the start of this paper I will explain what is meant by a Bayesian interpretation of probability, and the method through which Bayesian confirmation theory works. In doing this I will also introduce the notation used throughout. Proceeding from this I will go on to show the seeming failure of such a Bayesian interpretation when faced with the old evidence hypothesis, as introduced by Clark Glymour. For its clarification, I will split this problem into a selection of different cases, as used by Ellery Eells.

Following this, the argument I will show and explain for each case will be that they are based around a faulty definition of ‘confirming’ in Bayesian confirmation theory, and hence provide a corrected definition. Additionally, I then claim that critiques of these solutions are incorrect. I conclude by saying that while there are objections with the Bayesian theory, especially concerning the choice of prior probabilities, the old evidence hypothesis is no longer one of them.

In this paper we make the following observations about notation:

E: A piece of evidence supporting (or opposing) a hypothesis.

H: A hypothesis that is being tested based on some evidence.

$\text{Pr}(A)$: The probability that an event *A* occurs.

$\text{Pr}(A | B)$: A probability that an event *A* occurs, if *B* has already occurred.

1 What is Bayesian Confirmation Theory?

Bayesian confirmation theory has become one of the most predominant philosophical interpretations of probability after undergoing a revival at the end of the 20th century, especially in the scientific disciplines. Yet the ideas behind it extend back to Reverend

*Delivered at the BJUPS Summer Conference, 8 July 2017 at Wolfson College, University of Cambridge.

Thomas Bayes in the 1700s, and published posthumously by his friend Richard Price in 1763. The fundamental detail of Bayesian interpretation is based around conditional probabilities. Conditional probabilities tell us what the likelihood is that a theory is true, if some other evidence has already occurred. The formula showing this itself is incredibly simple:

$$\Pr(H | E) = \frac{\Pr(E | H) \Pr(H)}{\Pr(E)}$$

What this equation shows is the likelihood of H occurring, given that we know a piece of evidence E .

There are a few other useful forms of Bayes' theorem, and I will include the other form which will be of great use later on in this paper.

$$\Pr(H | E) = \frac{\Pr(H)}{\Pr(H) + \frac{\Pr(\neg H) \Pr(E | \neg H)}{\Pr(E | H)}}$$

Such equations appear to hold little philosophical significance when taken by themselves, however this is not necessarily the case. When we receive more information, we change our beliefs to fit with what we now know. Bayes' theorem does an excellent job of modelling such a belief change. By setting our probability of the hypothesis equal to $\Pr(H | E)$ after calculation, we have managed to achieve a probability for our new belief in the hypothesis. This is known as conditionalisation.

Conditionalisation easily gives us a way to measure whether or not a piece of evidence confirms a hypothesis, leading to what is known as Bayesian confirmation theory, which has the following hypothesis.

(Pt) Evidence E confirms a hypothesis H iff $\Pr(H) < \Pr(H | E)$

An example of this is given in the original paper, setting out Bayes' theorem (Bayes & Price 1763, 18–19). Imagine a man who has just emerged from a cave from the first time and sees the sun. However, after it goes down, he does not know if he will ever see it again. Yet, every day in the future he sees the sun rise; he then updates and increases the probability he assigns to his prediction that the sun will rise again the next day. All the time he gets more and more confident that the sun will always rise, until its probability is almost 1. It is by this principle that Bayesian confirmation theory works.

A point should be made here about subjective and objective probabilities in Bayesianism. This paper makes clear use of subjective Bayesianism and the idea that the prior probability assignment of a hypothesis is up to the choice of the actor who is choosing it. These are based on social factors experienced by the actor. Furthermore, such a choice of subjective probability leads to charges that this form of Bayesianism is not fit

for purpose. How could something so subjective and variable to an individual's own personal choice hope to explain science, something widely regarded for its objective nature?

While the arguments between subjective and objective Bayesianism are too numerous for this text, the argument that I make is one which seems to fit best into the nature of scientific investigation. At the beginning and in the early stages of a scientific investigation, individual scientists will obviously take varying, subjective, views on how likely different hypothesis will be. The expectation, given these circumstances, of finding a prior probability which is completely objective and that all the scientists agree on, is completely ludicrous and in the realm of fantasy. What should then happen is that, as evidence is received, scientific opinions will begin to converge, and while the idea of this will usually lead objective Bayesians (and those of other interpretations) to mimic Keynes and remark that in the long run we are all dead, this is not actually the case in this situation. In fact, with scientific hypothesis, opinions often tend to converge incredibly quickly after a relatively small amount of data. It is clear that Bayes theorem provides an excellent framework in order to model such an occurrence.

2 The Old Evidence Problem

What happens however, if you find evidence supporting a theory in advance of that theory ever even being introduced, as is not uncommon to do? In other words, old evidence supporting a new hypothesis. Surely that would then, intuitively, make the probability of the hypothesis being true, much more likely?

According to Bayesian epistemology, at first glance this would appear not to be true, which would make the Bayesian interpretation of probability utterly hopeless. This issue was first introduced by Clark Glymour, who stated

Scientists commonly argue for their theories from evidence known long before the theories were introduced. Copernicus argued for his theory using observations made over the course of millenia, not on the basis of any startling new predictions derived from the theory, and presumably it was on the basis of such arguments that he won the adherence of his early disciples. Newton argued for universal gravitation using Kepler's second and third laws, established before the *Principia* was published. The argument that Einstein gave in 1915 for his gravitational field equations was that they explained the anomalous advance of the perihelion of Mercury, established more than half a century earlier. Other physicists found the argument enormously forceful, and it is a fair conjecture that without it the British would not have mounted the famous eclipse

expedition of 1919. Old evidence can in fact confirm new theory, but according to Bayesian kinematics it cannot (Glymour 1980, 85–93).

Why is this so? Suppose E is a piece of evidence which is already known, and has been for some time, and suppose that H is a hypothesis. It is then later discovered that H predicts the evidence E . Using this we try to update our probability using Bayes' model.

$$\Pr(H | E) = \frac{\Pr(E | H) \Pr(H)}{\Pr(E)}$$

In this equation, $\Pr(E | H)$ is obviously equal to 1, as we have shown that H does imply E . As we also know, however, E does occur, so we must set $\Pr(E) = 1$. This obviously leads to

$$\Pr(H | E) = \Pr(H)$$

seemingly telling us that evidence which has been discovered in advance to the hypothesis cannot be used to change our belief in the truth of the hypothesis. This is obviously completely wrong, as is shown by the examples that Glymour cites.

3 A Clarification of the Old Evidence Problem

However, before solving this problem we must first clarify it. Eells argues that in fact the old evidence problem actually consists of a variety of different problems, which Garber claims can be separated between those that are ahistorical and those that are historical (Eells 1990, 207). These problems are as follows.

- (1) The problem of old new evidence: the hypothesis H is known before the evidence E ; however, $\Pr(E) = 1$ now.
- (2) The problem of old evidence these problems are those where the evidence E is known before the hypothesis H is known.
 - (a) The problem of old old evidence: occurs when H has been known for some time
 - (i) H was a hypothesis specifically designed to explain evidence E .
 - (ii) H was a hypothesis not specifically designed to explain the evidence E .

4 Solutions to the Problem

One major issue, which leads to people such as Glymour to misunderstand the Bayesian interpretation of probability, and thus to suggest perceived issues such as this one, is

that they appear to misinterpret hypothesis and evidence as being a two-way relation between evidence confirming hypothesis, in isolation from anything known previous. Yet, this is clearly not the case.

Bayesian confirmation theory is in fact, very clearly a three-way relationship between a piece of evidence, a hypothesis being tested, and some background knowledge K . That background knowledge being the set of all the previous evidence confirming or disconfirming the hypothesis prior to the current evidence being tested, and hence making up the current probability that we believe the hypothesis to be true.

We must bear in mind the key tenet to Bayesian confirmation theory, that once evidence has been tested, $\Pr(H) := \Pr(H, E)$, becoming the new prior probability for the next piece of evidence being tested. Therefore, the evidence E being true (or false) is added to the background knowledge, and now forms part of the probability that a hypothesis is true for testing the next piece of evidence.

This is obvious from example, such as this: Adam has been killed and John was seen near to the scene that it happened. Our hypothesis is that John killed Adam, and we test, using Bayes' theory, how much the probability of this hypothesis increases knowing that John was near the scene. We then make this new probability our probability that the hypothesis is true, and the knowledge that John was near the scene becomes a part of our background knowledge.

5 Solution to the Old New Evidence Problem

The understanding of such background knowledge immediately allows us to attempt to tackle the old new evidence problem, the problem in which we first formulated the hypothesis, and have subsequently discovered the evidence. Originally, the evidence does increase confidence in the hypothesis, however in the future $\Pr(E) = 1$, and can no longer increase confidence in H . This, surely E is still good evidence for the hypothesis though, and should still confirm it?

The solution to this follows very easily once you know of the background knowledge involved in Bayesian confirmation theory. One should first realise that the basis of this problem is on a misunderstanding of what the word confirm means in Bayesian confirmation theory, and when and what evidence confirms a hypothesis. To confirm in Bayesian confirmation theory, is to test that a piece of evidence is more likely to make a hypothesis true, and then changing the probability of the hypothesis to consider this. The evidence has then become learned.

Once an evidence has been learned, the evidence has already increased our probability that the hypothesis is correct, and we have updated the probability of the hypothesis to incorporate this. So following from this, that piece of evidence can no longer increase

the probability that the hypothesis is true, because it already has done exactly that. If it could, you could arbitrarily keep using the same piece of evidence to constantly increase the probability of the hypothesis.

Hence, going back to the example of a murder having taken place, we have performed the calculations that finding a knife in Johns apartment will increase the probability John killed Adam. We now make this new, increased probability the probability that John killed Adam. It is clear from this that simply knowing the existence of the knife cannot now increase confidence in H any further because our probability that John killed Adam is already based on us having the knowledge of the knife.

6 Solutions when the hypothesis has been designed originally to explain the evidence

Obvious solutions also exist for the old old evidence problem and the old new evidence problems in the cases where the hypothesis has been designed originally to explain the evidence. If the creation of the hypothesis has been to explain a piece of evidence, then that evidence itself is no longer evidence for the hypothesis. It is obvious in such a situation that our prior probability of the hypothesis is already based on knowing that it confirms the evidence, hence the evidence begins as part of our background knowledge. Therefore, to confirm that the hypothesis is actually good, we must find some new evidence, independent of the original evidence it was based on.

An example of this is as follows, firstly let us say that we find a knife covered in blood. We decide on the hypothesis that a murder has taken place, because that would explain the evidence of a knife covered in blood. We decide on a prior probability that such a murder has taken place. However, as we already know about a knife and have based out theory on that, our prior probability must begin by taking into account that we know about the knife. Hence such evidence can no longer increase our probability that the hypothesis is true, as it begins having already been learned, so it is already part of the background knowledge.

7 An Aside to Proposition 1

We stated in proposition 1 that

(Pt) Evidence E confirms a hypothesis H iff $\Pr(H) < \Pr(H | E)$

It has been made clear so far that such a definition does not accurately describe just how Bayesian confirmation theory uses evidence to confirm a hypothesis that is being

tested. Hence I would like to put forth a stricter proposition for evidence confirming a hypothesis.

(Pr') Evidence E confirms a hypothesis H iff $\Pr(H) < \Pr(H | E)$ relative to known background knowledge K .

Such a proposition neatly cleans up the previous two issues by making it clear why the evidence for the hypothesis does not seem to confirm said hypothesis, whilst more accurately working within the more commonly understood framework of Bayesian confirmation theory.

It clearly does this by noting that the same piece of evidence cannot confirm a hypothesis which it has already confirmed, in the context by which we understand the word confirm in Bayesian confirmation theory, which is to increase the probability that a hypothesis is true.

As will be shown, a solution to the rest of the old evidence problem follows from such a proposition, and is essentially the solution that Howson and Urbach gave for solving such a problem.

8 The Howson–Urbach Solution to The Old Evidence Hypothesis

The old evidence problem occurs in the situation when the evidence is known before the hypothesis, however sometime after the discovery of the hypothesis it is discovered that this evidence supports the hypothesis. Clearly in such a situation, $\Pr(E) = 1$.

Howson and Urbach provide a solution which is based around a similar idea to those involving background knowledge (Howson and Urbach 2006, 297–301). A piece of evidence does not show a change in the belief in the hypothesis, but instead shows a change in the belief in the hypothesis given the background knowledge.

The issue with the evidence E in this case, however, is that the evidence is already a part of the background knowledge K . This means that the evidence cannot itself create any change in the probability that the hypothesis is correct. We must instead remove the evidence from the background knowledge as much as we possibly can. We will term this new background knowledge $K - \{E\}$. Howson and Urbach themselves give an example of this (*ibid.*, 300–301). This is based around the observation of the anomalous precession of Mercury's perihelion, and how many years later it served as evidence for Einstein's general theory of relativity.

Let's use the following form of Bayes' theorem:

$$\Pr(H | E) = \frac{\Pr(H)}{\Pr(H) + \frac{\Pr(\neg H)\Pr(E|\neg H)}{\Pr(E|H)}}$$

We let H be the general theory of relativity and E be the data of Mercury's perihelion. In this, Pr refers to the probability, except it does not know E . As we have by assumption that $\text{Pr}(E | H) = 1$, so the formula becomes

$$\text{Pr}(H | E) = \frac{\text{Pr}(H)}{\text{Pr}(H) + \text{Pr}(\neg H) \text{Pr}(E | \neg H)}$$

As $\text{Pr}(E | \neg H)$ is the probability of E given anything except our hypothesis, so by expanding this out we end up with the sum $\sum \text{Pr}(E | H_i) \text{Pr}(H_i)$, where each H_i are the alternatives to the hypothesis H . In the current example, at the time the only alternative to the general theory of relativity was the classical gravitational theory, and so is clearly the only H .

As stated above, the function Pr does not know E so cannot take it into account. Hence, instead probabilities must be judged on all other background knowledge. As known, E is an anomalous result for the classical gravitational theory, and as such the probability of E given H_i is an incredibly small number we will term ε . So $\text{Pr}(E | H_i) = \varepsilon$.

What must next be evaluated is $\text{Pr}(H)$, once again against a set of background knowledge not including E . As E is not yet known to confirm H , $\text{Pr}(H)$ should be relatively small compared to the probability of the classical gravitational theory, $\text{Pr}(H_i)$. It follows from this, that $\text{Pr}(H | E) > \text{Pr}(H)$, thus showing that E , the old evidence, confirms H .

9 Glymour's Criticism

Glymour himself, anticipating such a solution, provided his own criticism to this answer to his problem (Glymour 1980 87). This criticism is that there is no such way to compute what $\text{Pr}(E)$ would be. For example, how could you compute a number telling you how much an effect the anomalous advance of the perihelion of Mercury had on confirming General Relativity?

The answer to this criticism is straight forward, however, if using a subjective Bayesian interpretation. This is something which Glymour and such critics seem to misunderstand. Whilst it may be difficult to comprehend what $\text{Pr}(H)$ would be, if it was assumed that you did not know E , however it is certainly possible, and will be based on the way that the person views the problem at hand, something completely acceptable in such subjective Bayesianism. Not only this, but there will be constraints in place due to the background knowledge, which will help to confirm the result.

10 Chihara's Criticism

Chihara puts forth a variety of issues which will be dealt with individually. These objections concern the nature of the set K of background knowledge and the removal of E from it. Chihara puts forth two separate concerns of the following form.

Problem 1 If the background knowledge K is supposed to always be a deductively closest set of sentences, then the set subtraction $K - \{E\}$ will not be well defined (that is, it would be impossible to classify all objects as either contained or not contained in $K - \{E\}$ but not both) if any sentences exist in K which lead to E . This also means that there is no unambiguous way to remove any such sentences and hence exists no uniform way to remove E (Chihara 1987, 553).

Problem 2 If we assume that K is deductively closed, and we find a way to just remove E from it, then any consequences of E will continue to remain in our new, relativised background knowledge. Information we would not have known had the experiments entailing E not taken place would still be included. This will mean that E could still be extremely probable. While this would not be as bad as assigning $\text{Pr}(E) = 1$, it would still be a serious issue (ibid., 553–554).

To first solve problem 1 we must first disregard the idea of the deductive closure of the background knowledge. The representation of one's background information is not some deductively closed set, and is in fact more akin to a set of logically discrete statements. While such a description of the background information is certainly an idealisation, it is still a fairly accurate way of describing it. The fact that the background knowledge is not deductively closed does not matter anyhow using Bayes theorem. A probability function relativised to a body of background information will assign the value of 1 to any possible deductive consequences of the sentences in the background knowledge. This does not matter if such a consequence itself was originally included in the background knowledge or not. We must also regard K as an axiomatisation of the background knowledge, and hence regard the difference of sets $K - \{E\}$ as the set of elements in K not including E .

The major issue with this problem is that such background knowledge can be axiomatised in various different, equivalent, ways. This means that the effect of removing a piece of information from different, equivalent axiomatisations can lead to non equivalent consequences. Hence there is no uniform method of removing information from the background knowledge. Howson lays out an example of this (Howson 1991, 550). $\{a, b\}$ is equivalent logically to $\{a, a \rightarrow b\}$, where a and b are logically independent sentences. If we wanted to remove a from this however, we would get two different sets of consequences, with b in the former but not the latter.

How do we solve such a problem then? Once again the solution is that of subjective Bayesianism. We must not think of the existence of a purely objective logic of induction which does not exist, and instead recognise that it is subjective. Such an issue is entirely consistent with this theory, as various people may view a problem in differing ways, and hence structure the information according to this. Thus, what people remove from the background knowledge will vary by how that person sees the problem. This allows people's estimation of the confirming evidence to vary. This can be combined with the ideas of Suzuki, who shows, using the AGM theory of belief revision, that a consistent probabilistic contraction function exists to represent deleting E from K , relative to some plausible boundary conditions on these functions (Suzuki 2005, 105–126). It is clear from this that problem 1, while looking taxing, is actually very solvable.

Problem 2 is probably the easiest to solve; in fact, it shouldn't be considered a problem at all. If the evidence E is very probable relative to $K - \{E\}$, because such background knowledge still contains consequences from the observations that gave rise to E , then obviously the evidence is simply not a strong confirmation of H . This is exactly how Bayesian confirmation theory is supposed to function. As long as we remove everything from K that is logically dependent on the evidence E , then the probability will be less than 1, hence the old evidence problem does not arise.

II Conclusion

In conclusion, it is clear that the so called old evidence problem to the Bayesian interpretation is not actually a problem at all. The supposed problem clearly arises both from a misunderstanding of both the concept of 'confirmation' in Bayesian confirmation theory, and the misguided assumption of the existence of a purely objective logic of induction. It is because of the first reason that a much stronger principle for Bayesian confirmation theory is needed to be provided, which now has been, taking into account that evidence can only confirm a hypothesis relative to some background knowledge. We have with the second reason however, strayed into the controversial topic of an objective or subjective Bayesian interpretation. While I make it clear that a subjective Bayesian interpretation is the only one which makes sense, it cannot be proved that no further constraints may exist to further restrict the freedom to which prior probabilities can be chosen.

References

- [1] Bayes, T. and R. Price (1763). *An Essay Towards solving a Problem in the Doctrine of Chances*.
- [2] Chihara, C. (1987) ‘Some Unresolved Problems in Bayesian Confirmation Theory’. *The British Journal for the Philosophy of Science* 38: 551–560.
- [3] Eells, E. (1990). ‘Bayesian Problems of Old Evidence’. C. W. Savage (ed.) *Scientific Theories*. Minneapolis : University of Minnesota Press.
- [4] Glymour, C. (1980). *Theory and Evidence*. Princeton: Princeton University Press.
- [5] Howson, C. (1991). ‘The “Old Evidence” Problem’. *The British Journal for the Philosophy of Science* 42: 547–555.
- [6] Howson, C. and P. Urbach. (2006). *Scientific Reasoning: The Bayesian Approach, 3rd Edition*. Chicago: The Open Court Publishing Company.
- [7] Suzuki, S. (2005). ‘The Old Evidence Problem and AGM Theory’. *Annals of the Japan Association for Philosophy of Science* 13: 105–126.

Foucaultian Interpretation of Sexuality and Its Positive Implications: An Account of Change

Anna Giordano

King's College London

Introduction

In the introductory speech for his module at the College De France, Foucault specifies, among other things, ‘the world is not the accomplice of our knowledge’ (1970, as quoted in Young 1981). In other words, the universe is not at immediate disposal; it is not *there* in an intelligible form. If we want our knowledge to adapt to the universe, what is required is the intense work of modifying our concepts and structures such that it takes on an intelligible form.

To take one interesting example, knowing and coming to understand one’s own sexuality represents, in Foucault’s view, much more than we may normally suppose. Arriving at an understanding of one’s sexuality implies acquiring three new meanings: first, the concept of *knowledge* as produced by discourse; secondly, a concept of the ‘self’, called by Foucault the *subject*; lastly, a concept of *sexuality*, derived from an analysis of its history. These three concepts are, however, not to be thought of as three independently existing objects in the world; rather, they are interconnected: to know one’s own sexuality, for Foucault, is something that requires a deep investigation, which is not just critical, but above all *genealogical*, and does not presuppose the existence of an absolute truth of sexuality divorced from historical and social context.

In my essay, I begin from the Foucauldian idea of truth, as produced by discourse and constituting knowledge, which circulates through relations of power. My aim will be to show that, for Foucault, knowing one’s own sexuality would mean nothing but acquiring and sharing the concepts (and eventually recognizing oneself into a category) that power-knowledge happened to create, in the given historical moment and context in which one happens to live as the subject.

1 A brief overview of Foucault

In *Truth & Power*, Foucault states:

Truth is a thing of this world: it is produced only by virtue of multiple forms of constraint. And it induces regular effects of power. Each society

has its regime of truth, its ‘general politics’ of truth: that is, the types of discourse which it accepts and makes function as true; the mechanisms and instances which enable one to distinguish true and false statements. (Foucault, as quoted in Rabinow 1991).

What is considered ‘true’, be it a scientific, political or social statement, is not true in virtue of a correspondence with some external fact; rather, when a speech-act, claim or utterance is said to be ‘true’, according to Foucault, it is so in virtue of the fact that a particular form of discourse has developed about this speech-act, and certain entrenched beliefs and actions circulate within a particular community. The ‘truth’ of a discourse constitutes knowledge, and proliferates throughout the community through relations, namely interactions in all their forms and social variable power-dynamics (e.g. teacher-student, parent-child, doctor-patient, government-citizen, etc.). Consequently, since truth is dictated by social structures, social power lies in these very relations.

To this extent, Foucault’s book, *The History of Sexuality* (1978), is a genealogical investigation of a number of concepts surrounding sexuality. Foucault aims to analyze the discourses around sex and sexuality since the sixteenth century. Foucault does so, not in order to give an account of what sexuality is *per se*, as if sexuality can be divorced from its history, but rather in order to understand which force-relations were at work in the development of these concepts, and thus which power-knowledge created the ‘truth’ of sexuality at different historical moments, and how they changed over time.

What does it mean, according to Foucault, to ‘know’ something? We must not *assume* that truth-as-correspondence is the correct way to view our relationship with the world; rather, we must investigate how we use the word, and understand that the truth of something is produced by discourse, which come into existence around an object of communal inquiry. It follows that ‘knowing’ something simply means to come into contact with certain discourses raised at a given historical time. How can those discourses produce truth? It is this question that Foucault attempts to answer. To this end, he retraces the relations of power and institutions in their several forms, which at different moments throughout history have been able to make some things talked about and some others ignored. In this account, the existence of taboos is not seen as the prohibition of a power that comes from above and establishes what can be said and what must be hidden. On the contrary, taboos are themselves part of the structure of the discourse able to produce the truth about something. Power/knowledge is thus a notion that refers to the fact that power is not something that is possessed and conquered; power loses in Foucault its usual negative characterization. It becomes something that comes from everywhere and spreads everywhere, because the means through which it moves are knowledge and relations. Power, through the force relations, which it is made up of, creates discourse about things and thus knowledge about stuff that is then put into circulation by the same individuals, who constitute those

power-relations (1978). It is through this process that knowledge proliferates. Following this line of thought, it seems undeniable that we cannot really disentangle power and knowledge: the former, through the relations it involves, creates the second by using discursive technique; the latter proliferates and makes those relations keep the creative power they hold.

2 Sex and sexuality

I will now focus on the concepts of sex and sexuality. The Foucauldian investigation opens with the repressive hypothesis, which is a common view about sex, according to which during the sixteenth century discourses about sex were repressed in all their forms, until the emergence of Freudian theories, which were intimately based on that repression and considered as a liberation from it (1978). Foucault observes that the opposite is the case: the practice of confession was developed as a novel form of discourse about sex. Although sex was not talked about in the domestic sphere, the Church adopts certain social roles to produce the truth about sex. This rules out the hypothesis that sex was hidden for capitalistic reasons, namely that sex was concealed because the society was concerned with the production of labor only, and thus sex was just a means of reproduction, and pleasures were unnecessary. The ongoing talking about sex, in order to free individuals from its alleged repression, resulted in a transposition of the practice of confession from the religious sphere to all the other fields of life. While other societies came to develop an *ars erotica*, Western society did not, by internalizing the practice of confession in science. This produced a *scientia sexualis*, which purported to find the essential, theoretically-neutral truth about sex, without realizing that it was indeed producing the truth about sex in the same way as the Church centuries previously.

As Foucault notes, ‘Situated at the point of intersection of a technique of confession and a scientific discursivity’ the concept of sexuality itself emerges at this point. Resulting as a product of the new *scientia sexualis*, sexuality was ‘defined as being ‘by nature’: a domain susceptible to pathological processes, and hence calling for therapeutic or normalizing interventions’ (Foucault 1978). Foucault proceeds indeed in speculating that the bourgeois society emerged in the 19th century, rather than refuting sex, ‘put into operation an entire machinery for producing true discourses concerning it’ (1978). The discourse around sex, far from being regulated by the law, is regulated by the norm: discourse, by creating and making knowledge about something proliferate, produces and itself depends on the norm. It establishes what is normal and detaches it from what it does not consider as such, so creating new categories and new sexualities.

Similarly, power has to be considered in its creative form, because of interactions of force relations and spreading and exerted through them. As Foucault puts it: ‘we must

at the same time conceive of sex without the law and power without the king' (1978). Thus, it is in virtue of the power relations at work in the sixteenth century that sexuality became an area of investigation. Foucault goes on analyzing the four figures of sexuality that emerge from the concern about sex during the eighteenth century: the 'hysterical woman', the 'masturbating child', the 'Malthusian couple' and the 'perverse adult'. He observes how these four figures were constituting the very concept of sexuality itself, which is then nothing but a historical construct. 'Sexuality' is not something that exists independently of our thinking about it: it is not a self-dependent reality, which tries to hide itself and thus needs to be discovered. Foucault defines it as 'a great surface network in which the stimulation of bodies, the intensification of pleasure, the incitement to discourse, the formation of special knowledges, the straitening of controls and resistances, are linked to one another, in accordance with a few major strategies of knowledge and power' (1978). This 'question' of sex resulted in two processes: on the one hand 'we demand sex to speak the truth' and on the other we want that sex itself reveals us our own truth. It is at this point that our constitution as subjects take place.

Power/knowledge, through discourses and relations, creates and perpetuates that *scientia sexualis* which defines the norms of sex. The individual who happens to live in such a context is thus led to internalize these norms, meanwhile monitoring him/herself to conform to these norms. Using Foucault's words, 'this form of power applies itself to immediate everyday life which categorizes the individual, marks him by his own individuality, attaches him to his own identity, imposes a law of truth on him which he must recognize and which others have to recognize in him. It is a form of power which makes individuals subjects' (Foucault 1994). Sexuality is thus an historical construct, which establishes the norms of sex and creates categories. By doing so, it leads one to adhere to those norms and recognize oneself in a category. Through this process, it is ultimately constituted a subject whose own identity is dictated from those norms and categories. As Taylor puts it, 'Foucault envisions the self as embodied, embedded in a social and cultural milieu, constituted by power relations, in short, thoroughly contextualized' (Taylor and Vintges 2004).

Many have focused on this new concept of the subject advanced by Foucault, in order to explain and explore current social issues. For this reason, Foucault has played a fundamental role in feminist theories. It has been argued, for example, that disciplinary power affect those who freely choose to subject themselves to the practices of self-surveillance and self-normalization. Bordo believes that in the light of this auto-subjection to disciplinary power, the perpetuation of the patriarchal system and the adaptation to standard of femininity are alimanted from the inside (1988). Following the same line of thought, Bartky (2010) believes that Foucault's analysis, by overlooking the forms of subjection that engender the feminine body, perpetuates sexism. The idea is that Foucault fails to recognize that the disciplinary practices, which apply to

men and women in order to create 'docile bodies', are more constraining for women than they are for men and produce 'a modality of embodiment that is peculiarly feminine'. Bartky divides them up into three categories: the first category includes those practices that aim to produce a body of a certain size and configuration; the second category refers to the practices which purport a social feminine standard of gestures, postures and movements; and the third category is about those practice that makes the female body an ornamental surface. Going back to Foucault's account of self-discipline, advanced in *Discipline & Punishment* (1995), and consisting in the idea that the inmates in prison – feeling surveilled – start to modify their behavior, Bartky retraces the same pattern in the behaviour of females in society. Males' eyes be the means of surveillance, feminine practices arise in order to meet the approval of males. Thus, women remain subordinated to men; the gender hierarchy is perpetuated.

That the disciplinary practices applying to women perpetuate an idea of femininity that is socially constructed, rather than natural, is undeniable. However, Bartky fails to recognize that masculinity works in just the same way. Reducing masculinity simply to 'the opposite of femininity', as Bartky puts it, is a mistake. Men, as well as women, face the same fear provoked in the inmates by the central guard of the Panoptic. Other men's and women's eyes represent that figure of surveillance, both for women and for men. To understand this, I want to focus on the example used by Bartky to highlight the differences between men and women when it comes to occupy a physical space in the world. A photographic report shows that women tend to occupy as less space as possible with their bodies. Men, on the contrary, appear freer and more comfortable: 'they sit with legs far apart and arms flung out at some distance from the body' (Bartky 2012). It is plain that for Bartky that of men is an enviable position. But it is necessary to shift perspective in order to see where the default is. The point is not who can sit in an objectively more comfortable position. Rather, we need to go further and understand that freedom consists in acting according to how one feels to act. That the standard of masculinity contemplates that men occupy lots of space with their bodies is constraining for those who would rather prefer to take little space and sit as a woman is expected to, but decide to adhere to the norm in order not to feel excluded or discriminated. In other words, one is not freer if one adapts oneself to a standard one feels not to belong to, despite the fact that that standard could appear as a privileged one from the outside.

I consider Foucault's analysis of sexuality in an optimistic way. It is true that the omnipresence of power/knowledge makes impossible for a subject to detach him/herself from it. However, we ought not to forget that each of us, for his/her being a subject, constitutes that power/knowledge too. This means that a subject has a chance to change the current sexualities or gendered power relations, from the inside. Using McNay's words, Foucault's view 'exposes the contingent and socially determined nature of sexuality and, thereby, frees the body from the regulatory fiction of heterosexuality

and opens up new realms in which bodily pleasures can be explored' (1992). In other words, if discourses are what constitute knowledge, then our words, gestures and practices matter to the extent that they can create new truths, new concepts about things.

Similarly, if power-relations involve each of us, then each subject has a role in, and is responsible for, the proliferation of new kinds of knowledge. Thus, despite the fact that Foucault does not draw any difference between disciplinary practices that apply to men and women, he nonetheless furnishes the means of investigation to carve into each of us. In *The Care of the Self*, he suggests the idea of 'knowing oneself' only by understanding who one really is, one can acknowledge that the standard is not inclusive enough, that the ideas of masculinity and femininity need to be challenged, that the categories do not suffice. This is the first step to change things: awareness.

Whether those social relations are based on disciplinary power only, or also on mutuality and reciprocity, is a source of debate. That the self is relational could mean two things. On one hand, it could be the case that a subject creates a concept of the self in virtue of communicative, reciprocal, mutual relations with others. And this would mean that 'mutual relations with others are both possible and crucial for how we think about the self' (Taylor & Vintges, 2004). This is the view of the self that feminists have tried to extrapolate from Foucault's late works in order to demonstrate that the care of the self needs to be social. As McWhorter asserts, 'putting care for one's self above all else might mean developing the competence and self-confidence that make true generosity possible, the peace of mind that makes real sharing possible, and the desire for community that makes honesty, patience and cooperation paramount' (McWhorter 1999).

On the other hand, the social relations Foucault talks about can simply be seen as normalizing and disciplinary, and this would deny their reciprocal and mutual dimension. On this view, the account of the self provided by Foucault does not constitute an appropriate basis for the development of feminist theories, because it does indeed lack a clear understanding of the mutual and reciprocal relations that are necessary for one to recognize herself in the others. As Taylor says, 'since the relations Foucault focuses on are relations of power, and in so far they are all strategic, this leaves no room for concept of mutuality and reciprocity others have seen in them' (2004).

I think the real merit of the Foucauldian investigation lies in the fact that it does not aim to say how things should be, but rather it enlightens how things are and how they can be transformed. For this reason, despite the fact that Amy's account of the relational-self, based on strategic power-relations only, is consistent with Foucault's works, nothing rules out the possibility that the subjects themselves can change this very concept of the self. In other words, if an account of mutuality and reciprocity embedded into relations is what is needed for improving self-identity and interactions with others, it is in our power to make room for those dimensions, despite the fact that

they are currently missing in our account of the self.

3 Conclusion

In conclusion, to know one's own sexuality would mean, for Foucault, several things. First, it would signify to acknowledge that knowledge does not coincide with the absolute truth, but rather it is the result of the discourses about different matters, which proliferate through the power relations. Second, it would mean to recognize that sexuality is not a mind-independent reality existing in the external world, but rather a concept created by the internalization of the practice of confession into science. It is then an historical construct whose truth is produced by power/knowledge relations, which establish the norm and create subjects. Each subject is such in virtue of his self-discipline and his attempts to conform to the norm and recognize himself into a category rather than another. Yet, despite the fact that this account seems limiting, I have outlined its positive implications. In virtue of Foucault's account of knowledge and power, being subject means having the possibility to change things and produce new truths about the world, simply by means of discursive practices, made spread through social relations.

References

- [1] Bartky, S.L. (2010). 'Foucault, Femininity, and the Modernization of Patriarchal Power', In *The Politics of Women's Bodies: Sexuality, Appearance, and Behavior*, 3rd ed. Rose Weitz (ed.). New York: Oxford University Press: 76–98.
- [2] Bordo, S. (1988). 'Anorexia Nervosa: Psychopathology as the Crystallization of Culture' in I. Diamond & L. Quinby (eds.) *Feminism and Foucault: Reflections on Resistance*, Boston: Northeastern University Press.
- [3] Foucault, M. (1978). *The History of Sexuality*, trans. R. Hurley, Penguin Books.
- [4] Foucault, M. (1988). *The Care Of The Self*. 1st ed. New York: Vintage Books.
- [5] Foucault, M. (1995). *Discipline And Punish*. 1st ed. New York: Vintage Books.
- [6] Young, R. (1981). 'Untying the Text: A Post-Structuralist Reader'. London and Boston: Routledge and Kegan Paul.
- [7] McNay, L. (1992). *Foucault and Feminism: Power, Gender and the Self*. Cambridge: Polity Press.

A Kuhnian Critique of the Feminist Movement

Martha Crowe

University of Aberdeen

Abstract

The aim of this essay is to utilise Kuhn's theories on the structure of scientific revolutions to examine the historical structure and future development of the feminist movement. I notice certain discrepancies in the ideology of modern feminism, especially 'second wave' feminism, specifically what appears to be two central premises. Furthermore, I apply Kuhn's theories to see whether there are any insights about progress within the feminist movement.

I first draw links between aspects of Kuhn's theory and the history of the feminist movement. From that point I focus on anomalies, aspects of the reigning paradigm in the second wave which have forced the revisions of the third.

I note one specific anomaly, within the concept of *essentialism*. I recognise a shift in feminist theory that relied on essentialism to one based on intersectionality. By focusing on these two waves of feminism I then project what I believe to be necessary developments in the movement, adaptations which are necessary in our changing and expanding world.

I conclude both theories show specific aspects of their respective waves of feminism and that the future of feminism may well follow a Kuhnian paradigmatic structure.

Introduction

Thomas Kuhn's essay, *The Structure of Scientific Revolutions*, dealt primarily with the history and development of one discipline: the natural sciences. He claimed that the development of science was not cumulative, but instead went through various stages or 'paradigms', which drastically altered which beliefs or practices were acceptable or unacceptable (Kuhn 1962). This, he argued, was in conflict with more classical conceptualisations of science, which viewed science as 'development by accumulation' of 'individual discoveries and inventions' (ibid., 2).

Since 1962, the phrase 'paradigm shifts' and the label 'paradigm' have been applied to numerous and various disciplines, movements and cultures in an attempt to describe social and cultural changes. By adopting Kuhn's approach, a number of fields now have a useful functional method through which standards and common bases for thought, as well as changes in shared beliefs within a community, can be explained.

As a social movement and academic discipline, feminism lends itself particularly well to Kuhn's theory of revolutions. This is due in part to a *prima facie* similarity between

the historically understood three ‘waves’ of feminism and the three stages of Kuhn’s classical paradigm-theory. In this essay I examine each metaphorical ‘wave’ of feminism with regard to the three stages Kuhn outlines in his essay: ‘The Route to Normal History, Normal History, and Crises’. Through the use of Kuhn’s model, I examine what he termed ‘anomalies’: aspects of the feminist movement which do not comply with the overarching paradigm and are thus ignored until their number and significance causes the paradigm to collapse (ibid.). At this point a new paradigm develops to account for the anomalies of the previous paradigm.

Specifically, I discuss one possible ‘anomaly’ within contemporary feminism. Essentialism and intersectionality exist in opposition as two defining currents of thought within the second and third wave, respectively. Essentialism was a key concept within the reigning paradigm built up within the second wave, and intersectionality came into being as a response to the unsatisfactory conclusions which essentialism prompted. The relationship between these two concepts is emblematic of both the relationship between the two waves and that of a reigning paradigm and its anomalies. In the history of feminism various anomalies existed and drove the emergence of third wave feminism. Thus, by interpreting the history and development of feminism using Kuhnian paradigm-theory, I examine the aspects of the movement which are particularly vital in contemporary feminism and in its future development.

Although Kuhn’s theories have been subject to many controversies since their initial formulation (Nickles 2009), these will not be the focus of my inquiry, but instead the usage of Kuhn’s theory as a functional model with which to examine the history of feminism as well as its current state. It is not my purpose to come to a conclusion as to the veracity of Kuhn’s theories. Instead my focus is upon the efficacy of Kuhn’s theory of paradigmatic revolutions in examining feminism and what this might indicate in terms of the development of feminism.

1 Thomas Kuhn’s Theory of Scientific Revolutions

In *The Structure of Scientific Revolutions* Thomas Kuhn opposed traditional views of the development of science as ‘the piecemeal process by which [...] items have been added, singly and in combination, to the ever growing stockpile that constitutes scientific knowledge’ (Kuhn 1962, 2). Instead he believed the natural sciences, both as a whole and in extremely specialised fields, were based upon a certain paradigmatic structure. A paradigm is a guiding force in further research as it prescribes a certain way of thinking about a discipline. Kuhn insists that a paradigm ‘goes beyond shared rules, assumptions and points of view’ as ‘paradigms can guide research even in the absence of rules’ (ibid., 42). This helps to shape and structure further inquiry and development of the subject. The science then continues, as a paradigm, until it is over-

turned due to an increasing number of discrepancies. Kuhn dubs these discrepancies 'anomalies'. The anomalies arise from 'the recognition that nature has violated the paradigm-induced expectations that govern normal science' (ibid., 53).

I propose that within the feminist movement a 'strong network of commitments – conceptual, theoretical, instrumental, and methodological' arose in its history, that the movement itself experienced a paradigmatic shift, and now is located at a point of crises (ibid., 42). The first stage consists of pre-paradigmatic schools of thought, which are based upon 'fact-gathering' and do not adhere to a unified theory. In the second stage, these 'facts' are consolidated into the paradigm, which shapes all further inquiry within the field. In the third stage the paradigm goes through a process of dissipation in which the acquired knowledge is recognised to be incompatible with certain facts accepted by the community. Accepting these accumulated incompatible facts causes a 'paradigm shift' to occur: the key laws and rules of the previous paradigm are altered or replaced to accommodate the overwhelming number anomalies which have arisen in the previous paradigm. A new school of thought must then account for these anomalies. The anomalies manifest themselves as issues or debates which are incompatible with the unifying ideology behind the second wave of feminism. Kuhn writes, 'Out-of-date theories are not in principle unscientific because they have been discarded' when writing about a change in paradigms (textitibid., 3). Similarly, feminist theory which is now highly critiqued and in many ways obsolete as a functional theory, i.e., essentialism, is not un-feminist. Indeed it is beneficial for the feminist movement as a whole and feminist philosophy to approach the changes and growth of feminism as not in opposition. Rather, feminists must examine the historical integrity of the viewpoints of their time, and recognise that feminism is also susceptible to shifts and alterations of what may appear to be indispensable aspects of its basic theories.

2 The First Wave: The Route to Normal Science

First wave feminism has entered history as being the time of the suffragettes, a period in which women truly began to form a group and collaborate in order to combat injustices within their society. Naturally of course, this focus on the suffragettes means that many influential women whom we would now refer to as 'feminist' were in effect removed from this first wave: '[E]arly fact-gathering is far more nearly random activity than the one that subsequent scientific development makes familiar', and it is also '...restricted to the wealth of data that lie ready to hand' (Kuhn 1962, 15-16).

As the object of inquiry is not the natural sciences but rather the feminist movement, there is a dissimilarity between the two: the fact-gathering in the natural sciences is broader in the history of feminism. This first stage within the feminist movement can be likened to 'experience-gathering'. By this phrase I mean to refer to the women that

developed from the personal experience of living under a patriarchal system the conceptual structures necessary to understand the power-imbalances in play, the women who wrote and spoke about these issues long before the existence of any academic feminism or the entrenchment of modes of thought, and who gathered together their own understanding of the inequalities that women were subject to and had to try and alleviate independently.

An examination of these first wave feminists shows that they had varying agendas, and focused upon the existence of ‘anomalies’ within their own existing paradigm – a patriarchal society which viewed women as inherently of lesser moral, social and intellectual worth than men. Miriam Brody writes, ‘[t]he mainstream feminism in (the UK and USA) after the ‘Vindication’ was concerned with achieving limited and specific reforms often for particular kinds of women, who, it was assumed, were without the conventional protection of the male sex’ (Brody 1992, 59). Little was done outside conceptions of the contemporary paradigm. Therefore, the points raised by women who were viewed as pioneers by the second wave differ greatly: the ‘fact-gathering’ was random, and include Mary Wollstonecraft’s *A Vindication Of The Rights Of Women*, Sojourner Truth’s oft-misquoted *Ain’t I A Woman?*, and Emily Davison’s act of martyrdom for women’s suffrage. Each of these women had a different approach to the issue, and due to the lack of an existing movement, their works were rejected by mainstream society and generally reviled. Two examples demonstrate the range of rejection: Davidson was repeatedly arrested and subject to forced feeding; Wollstonecraft was called a ‘hyena in petticoats’ by Horace Walpole in reaction to her *A Vindication Of The Rights Of Women* (Bentley 2005, para. 8).

Individually, these women acted and wrote of their thoughts and theories. Women gained the vote for the first time in multiple countries during a period of intense activity for the suffragettes. However, it is questionable as to whether this was purely due to activists’ efforts or other historical factors, such as World War One. Certainly, the suffragette movement and others working outside of the suffragettes did not lead to women’s emancipation, Wollstonecraft’s ‘Rights’ were not granted and suffragettes were viewed as outcasts of society. Feminism had experienced its beginnings through the work of these women, through their ‘fact-gathering’, yet the true paradigm was not yet formed.

3 The Second Wave: The Paradigm, Normal Science and Puzzle Solving

The second wave of feminism indicated the creation of a paradigm. Beginning in the 1960s and continuing throughout the 1970s, the ‘Women’s Liberation’ movement dealt with a large variety of issues (Greer 2000, 2). This second wave clearly defined what

was acceptable and unacceptable within the movement. It was accepted as given, for example, that women have been oppressed by men on a societal level, but assigned gender roles were not as fluid as in third wave feminism. The paradigm spawned a rich series of feminist writings, supported by the shared beliefs within the community. For example, members of the movement were able to write more openly on topics such as sexuality, and women's rights in the home and in the work place. Yet, second wave feminism was dependent upon underlying concepts and assumptions about what it meant to be a woman and a feminist. Essentialism was one such prominent theory within second wave feminism which was vital in providing a consistency of thought to the movement. It prescribed the view that one should identify as a woman above all other allegiances. Essentialism was traditionally defined in philosophy as: 'The doctrine that it is correct to distinguish between those properties of a thing, or kind of thing, that are essential to it, and those that are merely accidental. [...] Essentialism is used in feminist writing of the view that females (or males) have an essential nature (e.g. nurturing and caring versus being aggressive and selfish), as opposed to differing by a variety of accidental or contingent features brought about by social forces' (Blackburn 1994, 126). Not only did this legitimise and encourage women to unite behind this assumed commonality, it also separated and separated the state of being female from other states of identity, thus giving it a certain prominence. Essentialism was vital in creating a shared movement, justifying the importance of feminism to all women. The idea of a 'sisterhood' and 'womankind' were propagated and accepted as ideals (Johnson 2003, 25). However, this almost utopian idealism of a female revolution was not entirely straightforward, and has been the subject of much backlash in today's third wave (Aronson 2003, 905).

4 The Third Wave: Anomalies, Crises and Re-tooling

When examining contemporary, third wave feminism it becomes clear that it is in philosophical conflict with much of the tenets of the second wave. By examining these points of conflict, the anomalies which can no longer be ignored, one can come to some conclusions as to the response of feminism and its future development.

Interestingly, the defining aspect of each of the waves of feminism has often been indicated through the use specific names for each wave. Although we can discuss feminism as whole, referring to the first, second and third wave as a recognisable total, each wave has possessed specific designations: the Suffragette movement, Women's Liberation and Gender Equality respectively. These designations signify a specific focus and understanding of the feminism experienced at the time and reflects the breadth of the movement. Within the third wave, Women's Liberation and Feminism shifted to Gender Equality. The movement was not exclusively for women, rather it was about levelling the playing field for all genders. This shift was keeping with the theoretical aspect

of third wave feminism. However, Germaine Greer claims that not only are the terms gender Equality and Women's Liberation non-interchangeable, but that they are in philosophical opposition, and that the re-labelling of the movement has led to a restructuring of the essential beliefs and goals of feminism (Greer 2000). She states that women post-1970's 'were settling for equality. Liberation struggles are not about assimilation but about asserting difference, endowing that difference with dignity and prestige, and insisting on it as a condition of self-definition and self-determination' (ibid., 2).

Women's Liberation was very much in keeping with the essentialist ideals present in the second wave, however the name also portrays that which is incompatible with the third wave. The idea that women are liberated is taken for granted, sexism and inequality are no longer discussed using that terminology. Additionally, the aforementioned fluidity where the definition of a woman is concerned indicated the need for change. This shift signifies a Kuhnian 're-tooling' which was experienced as the movement grew and brought about changes to our society. 'Re-tooling is an extravagance to be reserved for the occasion that demands it. The significance of the crises is the indication they provide that an occasion for retooling has arrived' (Kuhn 1962, 76).

4.1 Essentialism and Intersectionality

This concept of essentialism, which was such a driving force behind the second wave of feminism, has become controversial in contemporary, third wave feminism. Greer claims it and feminism's 'force has been dissipated among a flotilla of other concerns, [...], gay liberation, black liberation, anti-pornography' (Greer 2000). This dissipation she refers to is a part of a major innovation, or anomaly within third wave feminism. The concept of essentialism has been generally dismissed as naive and exclusionary (Johnson 2003). Indeed, that a movement based upon a common idea of 'woman' functions both restrictively along with being unifying can also be explained in terms of paradigms. It is an example of the suppression of issues which a paradigm can enforce. As Kuhn (1962) terms it is 'an attempt to force nature into the pre-formed and relatively inflexible box that the paradigm supplies' (Kuhn 1962, 24). 'Nature', in this case, is the issue of now-named intersectional oppression, and the 'inflexible box' is representative of the concept of essentialism and its importance.

There has been a transformation in feminist theory for the recognition of multiple layers of oppression. Essentialism has been rebutted with the concept of intersectionality, and to a certain extent replaced in terms of the prominence awarded to it. As a term, 'Intersectionality' was introduced by Kimberlie Crenshaw (1989). The etymological background arose due to Crenshaw's use of metaphorical intersections in explaining the 'black woman's experience' (Crenshaw 1989, 44):

Consider an analogy to traffic in an intersection, coming and going in all four directions. Discrimination, like traffic through an intersection, may flow in one direction, and it may flow in another. If an accident happens in an intersection, it can be caused by cars traveling from any number of directions and, sometimes, from all of them. Similarly, if a Black women [sic] is harmed because she is in the intersection, her injury could result from sex discrimination or race discrimination

The experience of many women under an essentialist-orientated feminism was that of a movement which did not fully encompass their day-to-day interactions with sexism. Bell Hooks states in the acknowledgments her book, *Ain't I A Woman*, 'Eight years ago, when I first began research for this book, discussions of "Black Women and Feminism" or "Racism and Feminism" were uncommon. Friends and strangers were quick to question and ridicule my concern with the lot of black women' (Hooks 1982, i.). The recognition of the importance of a more diverse movement is typical of and necessary for our modern existence, and a natural reaction to a paradigm.

Although the term intersectionality came from concerns over racism and sexism intersecting, it can be applied to many issues and anomalies. One such anomaly is the emergence of multitudinous aspects and shapes of sexuality which are now present within our social frameworks, and how they have been dismissed and refuted by second wave feminists. These include the standardisation of terms such as 'Asexual' ('a person who has no sexual feelings or desires' (Oxford Dictionaries, 2013)) and 'Pansexual' ('not limited or inhibited in sexual choice with regard to gender or activity' (Oxford Dictionaries, 2013)) in discussions of gender.

Additionally, intersectionality and essentialism are particularly relevant in assessing the relationship between feminism and transsexuality. Transsexuality was excluded from the second wave feminist movement and even reviled as being misogynistic by some (Raymond 1994). However, in contemporary thought this view is both outmoded and 'transphobic'. Under essentialism however, transsexuality was something separate from feminist aims, and almost in opposition to them: 'The insistence that manmade women be accepted as women is the institutional expression of the mistaken conviction that women are defective males' (Greer 2000, 81). This and other comments by Germaine Greer concerning transsexuality have caused outrage amongst queer activist and pro-transgender feminists. From the viewpoint of third wave feminism, transsexuality is almost a deconstruction of gender – and a statement of how truly performative gender is in its nature. Judith Butler articulated this view in her book, *Gender Trouble*. In this she states that 'the illusion of a stably sexed body, core gender identity, and (hetero) sexual orientation is perpetuated through repeated, stylised bodily performances that are performative in the sense that they are productive of the fiction of a stable identity, orientation, and sexed body as prior to the gendered behaviour?' (Bettcher 2009, para. 5.1).

The ‘question’ of transexuality therefore illuminates the differences in perception of gender itself within the second and third waves. Gender changes from something stable, utterly crucial to one’s identity and to one’s role in feminism, to a fluid performance, something which must move beyond a binary and exists purely due to one’s own choice (Butler 1990).

Finally, the questioning of essentialist theories has caused a crisis within feminism, represented not merely by the debates I have outlined but in many more aspects of the movement. In accordance with Kuhn’s theory, the future of feminism will depend on its response to these crises.

5 The Future of Feminism: The Response to Crises

The recent backlash towards the Woman’s Liberation movement has left a divided movement, one which is more nuanced, and more open to difference. However, through this division feminism has experienced an important shift; from that of a communal movement, a uniting force and call for change, to a discipline, an academic filter. ‘The institutionalisation of women’s studies at some visible colleges and universities has made scholars forget the founding tenets of women’s studies as they slavishly attempt to recreate it in the image of traditional disciplines’ (Stacey 2000, 1189). My personal opinion is that feminism has been relegated to this position due to its current situation in the Kuhnian structure. Furthermore, it is important for the movement to experience this phase as only then will it be able to develop further. Kuhn (1962) states that ‘confronted with anomalies or with crises scientists take a different attitude toward existing paradigms, and the nature of their research changes accordingly. The proliferation of competing articulations, the willingness to try anything, the expression of explicit discontent, the recourse to philosophy and to debate over fundamentals, all these are symptoms of a transition from normal to extraordinary research’ (Kuhn 1962, 9). Currently feminism is debating its fundamentals, including but not limited to essentialism, in order to move beyond them.

As seen in this essay, much of mainstream feminist theory has until recently dealt almost exclusively with white, middle class women, and been considered from that perspective (Hooks 1982). I believe that the new, emerging paradigm of feminism must change in that regard. Already the concept of intersectionality has appeared. In the future, feminism must address the concerns of globalisation as an issue pertinent to equality of the sexes and the liberation of women. As the world become more and more connected, the prejudices which hold that non-western women are in some manner ‘more oppressed’ than liberated, western women are not only outdated, but narrow minded and insulting (Al-Hassan Golley 2004). One of many points of tension in feminism today is the debate over the Islamic belief that wearing a headscarf is nec-

essay for women in order to follow their religion. Unfortunately, in France in 2011 the incorrectly-dubbed 'Burqa Ban' was put into place, depriving Muslim women the freedom to express their religion. Feminist ideology was used as a justification for this law, as some claimed that it liberated women who had been forced to wear a headscarf (Ramdani 2013). Future feminism should not be able to be perverted in such a manner, and must be adapted to show clearly what kind of a movement it will be, acknowledging the current state of our world. I eagerly anticipate the transition from a neutered, purely academic feminism to its next paradigm, as only then can it once again become a real force for change.

Conclusion

In conclusion, the structuring of feminism with regard to Kuhn's theories of paradigmatic revolutions has shown that the development of the movement does to some extent follow a recognisable pattern. Essentialism and intersectionality have proven themselves to be fascinating criteria with which to assess the transition feminism has made from the second to the third wave, and has highlighted feminism's reliance on a paradigm in order to be an effective force in the world. Despite this, the process of aligning a social movement with a philosophical theory concerning science requires a certain amount of generalisation. In order to construct my theory I found it difficult to thoroughly encompass the first wave of feminism, and was forced to equivalent a large section of history to the pre-paradigm stage. The question of whether I could truly represent such a large and multifaceted movement within 4000 words plagued me in my research. Despite this, I recognise that the modelling of feminism on Kuhn's theories did indeed lead me to draw conclusions about the future of feminism as well as questioning what feminism truly is, beyond paradigmatic trends.

References

- [1] Al-Hassan Golley, N. (2004). 'Is Feminism Relevant to Arab Women?', in *Third World Quarterly*, 2(3): 521–536.
- [2] Aronson, P. (2003). 'Feminists or "Postfeminists"?: Young Women's Attitudes toward Feminism and Gender Relations', in *Gender and Society*, 17(6), 903–922.
- [3] Asexual. (2013). In *Oxford Dictionaries*. Retrieved October 1, 2013, from <http://www.oxforddictionaries.com/>
- [4] Blackburn, S. (1994). 'Essentialism', in *Oxford Dictionary of Philosophy*. Oxford University Press, USA: 126.
- [5] Bentley, T. (2005). "Vindication": Mary Wollstonecraft's Sense and Sensibility', in *The New York Times* [Online] Available at http://www.nytimes.com/2005/05/29/books/review/vindication-mary-wollstonecrafts-sense-and-sensibility.html?_r=0
- [6] Bettcher, T. (2009). 'Feminist Perspectives on Trans Issue'. In E. N. Zalta (ed.), *Stanford Encyclopedia of Philosophy* [Online] Available at <https://plato.stanford.edu/entries/feminism-trans/>
- [7] Brody, M. and M. Wollstonecraft (1992) Introduction: 'Posterity Will, However, Be More Than Just'. *Feminism After The Vindication*, in M. Brody (ed.), *A Vindication Of The Rights Of Woman*. St Ives, England: Penguin Classics.
- [8] Butler, J. (1990). *Gender Trouble: Feminism And The Subversion Of Identity*. London, UK: Routledge.
- [9] Crenshaw, K. (1989). 'Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics', in *University of Chicago Legal Forum*: 139–167. [Online] Available at <http://politicalscience.tamu.edu/documents/faculty/Crenshaw-Demarginalizing.pdf>.
- [10] Greer, G. (2000) *The Whole Woman*. St Ives, England: Anchor.
- [11] Hooks, B. (1982) *Ain't I A Woman: Black Women And Feminism*. London, UK: Pluto Press.
- [12] Kuhn, T. (1962) *The Structure Of Scientific Revolutions*. (4th ed.) USA: University of Chicago Press.
- [13] Johnson, A. (2003) 'Seeking Sisterhood: A Voice From The Third Wave', in *Off Our Backs*, 33(9/10): 22–25.

- [14] Nickles, T. (2009). 'Scientific Revolutions'. in E.N. Zalta (ed.), *Stanford Encyclopedia of Philosophy* [Online] Available at <https://plato.stanford.edu/entries/scientific-revolutions/>
- [15] Pansexual. (2013). In *Oxford Dictionaries*. Retrieved October 1, 2013, from <http://www.oxforddictionaries.com/>
- [16] Ramdani, M. (2013). 'Don't follow France's burqa ban. It has curbed liberty and justice', in *The Guardian*, September 21, 2013.
- [17] Raymond, J.G. (1994). *The Transsexual Empire: The Making of The She-Male*. New York, NY: Teachers College Press.
- [18] Stacey, J. (2000). 'Is Academic Feminism an Oxymoron?', *Signs*, 25(4): 1189–1194.

MA Philosophy Durham University



MA Philosophy

Course structure

This one year programme (two years part-time) comprises taught modules and a substantial dissertation. Modules are taught via group seminars and one-to-one tutorials. Alongside core compulsory modules you can choose between a range of topic modules allowing you to develop your own research focus throughout the course. Examples of topic modules include:

- Current Issues in Ethics
- Philosophy of the Social Sciences
- Mind and Action
- Philosophical Issues in Science and Medicine
- Current Issues in Aesthetics and Theory of Art
- Current Issues in Metaphysics
- Phenomenology and the Sciences of Mind
- Gender Theory and Feminist Philosophy
- Ancient Philosophers on Necessity, Fate and Free Will
- Environmental Philosophy
- Business Ethics

Modules offered are subject to change.

For up-to-date information, see www.durham.ac.uk/philosophy/postgrad/taught/taughtphil

Career Opportunities

This MA course provides students with a wide number of important skills, including sophisticated analytic skills and intellectual clarity. This kind of study involves a high level of commitment, motivation and discipline, inherently employable qualities. It also provides an ideal academic environment for those who would like to study the subject at a higher level in preparation for a PhD.

MA students can benefit from a range of other activities in the department, including the department's postgraduate philosophy society (EIDOS), weekly research seminars and reading groups, and occasional conferences, workshops and Royal Institute of Philosophy lectures.

Entry requirements

A typical 2:1 classification or higher at undergraduate level or equivalent qualification with a substantial philosophy component.

If candidates wish to take the MA course with a research focus on Science, Medicine and Society, they may be considered if they have an undergraduate degree or equivalent qualification with another appropriate component, for example science-related subjects. For details, see www.durham.ac.uk/philosophy/postgrad/taught/taughtphil/resfocusscimedsci

For English language requirements, please see www.durham.ac.uk/philosophy/postgrad/taught/taughtphil

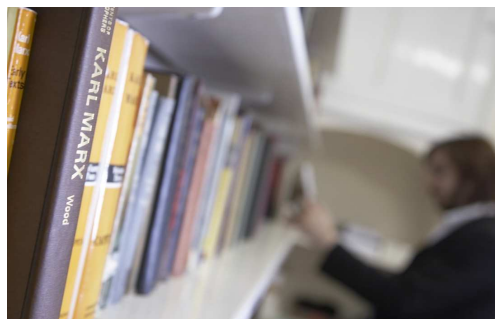
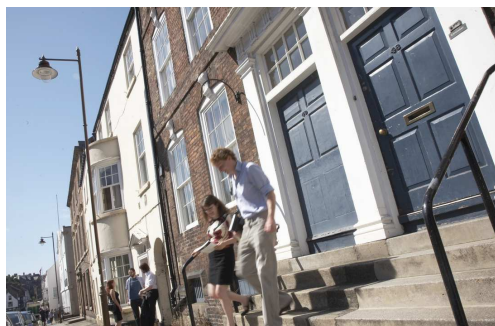
How to apply

You can apply online at www.durham.ac.uk/postgraduate/apply

Enquiries

For postgraduate enquiries, please contact: +44 (0)191 334 6553 or philosophy.pgsec@durham.ac.uk

All information provided was correct at the time of print, but is subject to change - for the latest information, please see www.durham.ac.uk/philosophy/postgrad/taught/taughtphil



Durham University and Durham University logo are registered Trade Marks of the University of Durham. Unless otherwise stated, all material in this publication is copyright of the University of Durham. The University makes every effort to ensure that the information contained here is accurate. Please note that the University's website is the most up to date source of information and we strongly recommend that you always visit the website before making any commitments.

Issue II (2)
Summer Conference 2017

A Dilemma for Global Egalitarianism
Peter Wilson

*Would a Life Without Death Be
Intolerable?*
Joshua Goh

*Predicate Reference and Metalinguistic
Nominalism*
Beatriz Santos

*Solutions to the Old Evidence Problem in
Bayesian Confirmation Theory*
Matthew Coates

*Does Sarah Stroud's Account of Epistemic
Partiality Conflict With Evidentialism?*
Daniella Meehan

*Foucaultian Interpretation of Sexuality
and Its Positive Implications*
Anna Giordano

A Marxist Review of Justice
Giulia Carlini

A Kuhnian Critique of the Feminist Movement
Martha Crowe

