Aporia

Aporia

Undergraduate Journal of the St Andrews Philosophy Society
Volume XX

Letter from the Editor

Dear Reader,

Our 20th issue marks a landmark achievement of the Aporia Journal of Philosophy.

As we celebrate and look back on how much this journal has grown, now for over a decade, this is an excellent opportunity to look back and reflect on what this project has achieved. This journal has provided a space for outstanding undergraduate philosophy. Sadly, this type of work is often overlooked. The reasons range from impostor syndrome and under-confident students, to overwhelming competition and gatekeeping in professional journals. We have an important role to play. We believe that there are excellent pieces that should be shared with the world to spark discussions and inform our views, whether or not the author is considered a professional philosopher.

This year we have seen unprecedented reach and interest in our journal. With a record number of submissions, it is fair to say this has been a highly competitive process. We have also had the largest editorial team to date. With over 10 reviewers (students of undergraduate and postgraduate level from around the world) and with 11 St Andrews undergraduate editors selected after a competitive interviewing process, Aporia has become a successful philosophical enterprise, much bigger than anyone could have predicted.

What a year it has been.

Brexit appears to be reaching a conclusion and the UK having a concrete deadline to leave the EU after years of uncertainty. The world has been turned on its head. An epidemic that sheds light on systemic problems and makes us question the meaning of work, society and life – familiar questions in the history of this discipline. The fire of BAME movements was reignited around the world with a unique strength that we hope will persist and guide us in improving our world and preventing past mistakes. It is only fitting that this volume includes a paper which discusses minorities and presents an example of deeply prejudiced beliefs about Indigenous Peoples.

The Aporia Journal prides itself in its diversity (e.g. our efforts to reach gender balance in the editorial team and strict anonymity for all authors until all submissions have been fully reviewed). In spite of this, we think we can, and should, do more. Not only would we like to see more submissions from minorities but we must do more to make Aporia an inclusive journal that welcomes the best philosophy. Everything else should be put aside.

On a personal level, publishing this issue marks the conclusion of my undergraduate studies (after two years of being part of the Aporia Executive Team and

four years of being involved in our remarkable philosophy society). I will miss St Andrews. I will miss the wonder, the determination and energy to learn, to philosophise and to think about how to improve our world from our own little corner. I will miss the amazing communities, the magical town and the view of the sea after leaving the seminar room.

I hope that you enjoy this issue as much as we have.

Sincerely,

Arvo Muñoz Morán

Acknowledgements

Editors

Social Media and Outreach

Reviewers

Ludovic Glasson.

María Josefina Villanueva

Philip Lawrence Sink

Rahul Prakash

Spyridon Galatis

Victor Karl Magnússon

Vsevolod Khusid

Contents

A Panpsychist Interpretation of Anne Conway's Metaphysics

ANDREW FYFFE*
University of St Andrews

Abstract   This paper proposes a panpsychist interpretation of Anne Conway's (1631-1679) metaphysics, as elucidated in *The Principles of the Most Ancient and Modern Philosophy*. Contemporary versions of panpsychism attempt to explain how consciousness is realised in the natural world. They posit that matter is intrinsically experiential, such that when it is arranged into the form of a human brain, it gives rise to human consciousness. Similarly, Conway argues that substance is constituted by both Body and Spirit. The former serves as an explanation of a substance's material properties, whereas the latter explains how a substance can have various kinds of perceptual experiences, as well as experiencing sensation and emotion. I argue that Conway uses her concept of Spirit to refer to the same set of experiential properties as our contemporary concept of consciousness does. Understood thus, Conway's metaphysical framework appears to embrace a form of panpsychism.

1   Introduction

Scholarly interest in Anne Conway (1631-1679) has often directed attention to her arguments in favour of 'vitalism'.[1] By virtue of this, philosophers have failed to notice or, at the very least, properly account for the panpsychist elements of her only extant work, *The Principles of the Most Ancient and Modern Philosophy*.

---

*Andrew Fyffe is a soon-to-be graduate of the University of St Andrews. Unable to pry himself from rural Fife, he will return to the department in September to undertake the MLitt in Philosophy. His primary interests lie in the so-called 'hard problem' of consciousness. In particular, he wishes to encourage scientific and philosophical investigations into consciousness which view it as a fundamental constituent of the world around us. He is also interested in conceptual engineering, normative reasons, and linguistic communication.

1. For example, see: Carolyn Merchant, "The Vitalism of Anne Conway: Its Impact on Leibniz' Concept of the Monad," *Journal of the History of Philosophy* 17 (1979): 255-269.

I have two aims in this essay: one interpretative and one normative. First, I propose an interpretation of Conway in which her concept of Spirit is understood as coextensive with our contemporary concept of consciousness. This will consist of two sections: (i) an exposition of Conway's metaphysics of substance, and (ii) an attempt to show that 'Spirit' and 'consciousness' refer to the same set of mental properties. Second, I argue that one should adopt my interpretation, as doing so highlights the definite correlations between Conway's metaphysics and that of contemporary panpsychists. This will consist of two sections: (i) an overview of contemporary panpsychism, and (ii) an attempt to situate Conway's views amidst current discussions of panpsychism.

2   Introducing: The Metaphysics of Anne Conway

In *The Principles of the Most Ancient and Modern Philosophy*, Anne Conway provides grounds for rejecting Cartesian dualism and Hobbesian materialism, whilst outlining a metaphysics which inherits the virtues of both. In this sense, Conway provides an intermediate between two philosophical extremes: viz. the reduction of the mental to the physical (à la Hobbes) and the ontological separation thereof (à la Descartes).

2.1   Anne Conway's Metaphysics of Substance

Conway's ontology is monistic insofar as it commits her to the existence of a single type of substance.[2] Despite her adherence to monism, Conway respects the Cartesian intuition that there is *some* distinction to be drawn between the mental and the physical.[3] Unlike her Cartesian contemporaries, however, Conway argues that there is no *essential* distinction between *that substance which possesses physical properties* and *that substance which possesses mental properties*.[4]

According to her framework, a substance can instantiate properties from either class, as substance is a coalescence of the physical (Body) and the mental (Spirit).[5] For her, the existence of mental and physical properties does not imply the existence of ontologically distinct mental and physical substances. Rather, it implies that mentality and corporeality are two modes of a single substance.[6]

---

2. Peter Lopston, "Introduction," in *The Principles of the Most Ancient and Modern Philosophy* (Martinus Nijhoff: London, 1982), 21.

3. The reader should note that I use 'physical' and 'material' interchangeably. Thus, physical and material properties are the same class of properties.

4. An essential distinction is one of essence or nature, e.g. Descartes' distinction between mental and physical substance.

5. Anne Conway, *The Principles of the Most Ancient and Modern Philosophy* (London: Martinus Nijhoff, 1982), 191.

6. This allows a substance to possess both mental and physical properties, as its nature is not restricted to the instantiation of either/or.

2.2   Conway's Concept of Spirit

There is nothing remarkable about Conway's discussion of Body. It shall suffice to say that her analysis mirrors Descartes' analysis of material substance. Simply put, Body is that which constitutes the physical aspects of a substance. That is, the properties of having a certain shape, breadth, and weight.[7] In this sense, Body is sufficient to metaphysically explain the physical aspects of an entity.

In Chapter VI of *The Principles*, Spirit is introduced to refer to that aspect of substance which allows for the instantiation of mental properties.[8] Spirit is, therefore, invoked to explain how, *qua* physical entity, a substance can have experiences in the visual, auditory, tactile, gustatory and olfactory modalities; sensation and emotion.[9]

Conway situates Spirit alongside Body as a fundamental part of the natural world. It constitutes one aspect of a substance's nature and is instantiated by all entities to varying degrees, ranging from rocks to God. As such, Conway presents us with a picture of the natural world exhibiting a hierarchy of mentality; with the most Spiritual substances occupying the highest and the most Bodily occupying the lowest echelons.[10] Most importantly, however, is there appears to be 'a scale of gradual shading' from the top to the bottom. Such that, regardless of the tier they occupy, all created substance is both mental and physical.[11]

2.3   The Interconvertible Nature of Spirit and Body

Another curious feature of Conway's metaphysics is the convertibility of Body and Spirit. God, *qua* infinitely Spiritual substance, has the power to alter the nature of particular substances. That is to say that God is responsible for conferring greater or lesser degrees of Body and Spirit onto each individual substance. And because Conway equates Spirit with perfection (i.e. Godliness), the more a substance ameliorates itself (e.g. morally), the greater the degree of Spirit God allows it to possess. Consequently, the further a substance moves away from God, the more corporeal it becomes.[12]

It is in this sense that created substances can be transformed (or can evolve) into different species of substance; realise greater or lesser degrees of mentality; and move further up the Spiritual hierarchy.[13] Such that, by divine contrivance, dust can become plants, plants can become dogs, dogs can become chimpanzees, and chimpanzees can become humans.[14]

---

7. Conway, *The Principles of the Most Ancient and Modern Philosophy*, 224-5.
8. *Ibid*, 180-81.
9. *Ibid*.
10. Lopston, "Introduction," p.15.
11. *Ibid*.
12. *Ibid*, 21.
13. *Ibid*.
14. *Ibid*, p.23.

2.4   The Key Features of Conway's Metaphysics

Firstly, Conway argues that there is only one type of substance, constituted by Body as well as Spirit. Secondly, where Body accounts for material properties, Spirit accounts for mental properties. Thirdly, Spirit is instantiated by *all* substances to varying degrees, thereby reflecting the difference in mental complexity across the natural world. Fourthly, through the will of God and in accordance with their moral conduct, substances can become more mentally complex by virtue of how much Spirit they possess.

3   The Coextension of 'Spirit' and 'Consciousness'

I will now proceed to show that Anne Conway's concept of Spirit and our contemporary concept of consciousness are coextensive. In other words, they are both used to denote the same set of mental properties.

3.1   Introducing: Phenomenal Consciousness

'Consciousness' herein refers to phenomenal consciousness.[15] Simply put, phenomenal consciousness is experience. To say that an entity is phenomenally conscious is to say that there is *something that it is like* to be that entity; that it has subjective experience. In this sense, phenomenal properties are experiential properties.[16]

Conscious states are a class of mental phenomena such as *seeing the colour blue* which have a distinct subjective feel. What it is like to be in a conscious state (that is, the subjective feel of the state) is determined by the set of experiential properties constitutive of that state.[17] For instance, feeling a sharp jolt of pain in your left leg feels a certain way; similar in nature yet phenomenally distinct from having toothache. On my account, conscious states *just are* experiential states. These states are individuated by their associated experiential properties. The totality of which determines what it is like for a subject to be in that state.

3.2   Taxonomy of Experiential Properties

From the recent literature, we can provide a coarse-grained taxonomy of experiential properties, including:

---

15. 'Consciousness' and 'experientiality' (and their associated properties/states) denote the same phenomena.

16. Ned Block, "On a Confusion About a Function of Consciousness," *Behavioural and Brain Sciences* 18 (1995): 230-31.

17. Thomas Nagel, "What is it Like to be a Bat?" in *Mortal Questions* (Cambridge University Press, 1979b), 166.

1. The various kinds of perceptual experience; such as *seeing* a red flower, *hearing* Frank Zappa's "Peaches en Regalia", *touching* a soft surface; and so on.

2. Bodily sensation; such as feeling dehydrated or cold.

3. Feelings of emotion; such as love, fear, desire, and regret.

4. Moods; such as happiness, sadness or boredom.[18] [19]

3.3 Textual Evidence for Coextension

Conway criticises Hobbes for thinking that an analysis of substance is exhausted by an analysis of extension. In doing so, one reduces material entities to 'mere Fabrick or dead Matter'.[20] Something was amiss in Hobbesian materialism – the absence of which rendered substance unfeeling and unthinking. And, although Conway does not use the term 'consciousness' in the *Principles*, it is clear that her concept of Spirit is referring to that class of experiential properties which Hobbesian materialism fails to account for.

For example, in Chapter IX of *The Principles*, Conway claims Spirit is that which gives substance the capacity for: 'Feeling, Sense, and Knowledge, Love, and Joy'.[21] In other words, the instantiation of Spirit allows for – or, at the very least, provides the potential for - a substance to undergo certain perceptual experiences ('Sense'), sensations ('Feeling'), emotions ('Love'), and moods ('Joy').[22]

3.4 Why Accept My Interpretation?

It appears that the properties of Spirit are experiential in the same sense that the properties of consciousness are experiential. States of love and joy, feelings and sensations, are all experiential states instantiated by conscious entities. And it is only by virtue of possessing Spirit that a substance can undergo these states. Conway therefore appears to be developing a theory concerned with the same mental phenomena as contemporary theories of consciousness.

---

18. David Chalmers, *The Conscious Mind* (Oxford University Press, 1996), 1-6. Michael Tye, "Philosophical Problems of Consciousness," in *Blackwell Companion to Consciousness*, ed. Velmans et al. (Oxford: Blackwell, 2006), 23.

19. This list is amenable to the narrow interpretative focus of this paper, not to substantive discussions of consciousness. As such, I do not take this taxonomy to be exhaustive. Thanks to the anonymous reviewer who emphasised this point.

20. Conway, *The Principles of the Most Ancient and Modern Philosophy*, 180.

21. *Ibid*, p.225.

22. One sympathetic to the so-called 'knowledge argument' against materialism may wish to correlate Conway's 'Knowledge' with the type of phenomenal knowledge discussed in: Frank Jackson, "Epiphenomenal Qualia."

Perhaps this common *explanandum* has eschewed scholarly acknowledgement due to Conway's use of 'Spirit' in lieu of 'consciousness'. Nevertheless, that Conway was concerned with consciousness is noted by her Early Modern contemporary and philosophic admirer, Gottfried Wilhelm Leibniz (1646-1716). In a letter dated 1697, Leibniz writes: 'My philosophical views approach somewhat closely those of the late Countess of Conway [. . .] because I hold that [. . .] everything takes place according to a living principle and according to final causes - *all things are full of life and consciousness.*'[23]

Therefore, given the sufficient textual evidence to substantiate my interpretation and the distinct lack of attention paid to the preceding considerations, one *should* read 'Spirit' as 'consciousness'. What's more, such an interpretation illuminates another underappreciated facet of Conway's metaphysics. Namely, the correlations it has with contemporary panpsychist theories of consciousness. Thus interpreted, Conway is understood as arguing for the view, as Leibniz put it, that all things are full of consciousness.

4   Introducing: Contemporary Panpsychism

Much like Conway's monism, contemporary panpsychism can be seen as an intermediary between reductive materialism and dualism. In fine, it is an attempt to explain how consciousness is realised in the natural world, whilst assuming the phenomenon to be irreducibly mental and ubiquitous throughout nature.[24]

4.1   The Intrinsic Nature of Matter

The general panpsychist commitment is that the basic constituents of the physical world such as atoms and quarks possess experiential properties as well as physical properties.[25] Matter, according to panpsychism, is *intrinsically* mental. Thus, when it is arranged into different kinds of organisms with different kinds of neurological structures, the basic experiential properties combine to realise different kinds of consciousness.[26] For example, when these basic constituents are arranged in the form of a human brain, the combination of their experiential properties gives rise to human consciousness. Panpsychism therefore posits that: (i) all physical entities possess some degree of consciousness by virtue of matter's intrinsic experientiality, and (ii) this explains how consciousness is realised in the physical world.

---

23. Gottfried Wilhelm Leibniz, letter to Thomas Burnett, 1697, in *Philosophischen Schriften*, ed. Gerhardt (Hildesheim: Olms, 1960). [Italics are my own.]

24. William Seager, "Panpsychism," in *The Oxford Handbook of Philosophy of Mind*, ed. McLaughlin et al. (Oxford University Press, 2009), 207.

25. Thomas Nagel, 'Panpsychism', in *Mortal Questions* (Cambridge University Press, 1979a), 181.

26. *Ibid*, 182.

4.2 Degrees of Experientiality

It must be stressed, however, that panpsychists are in no way committed to the claim that *all* substances instantiate full-blown phenomenal consciousness.[27] A panpsychist needn't claim, for example, that a rock possesses the hallmark features of mentality. As Chalmers notes, resistance to panpsychism tends to arise from a tacit conflation of experiential properties with other features of mentality. Most of which require a greater degree of material complexity for their instantiation.[28]

To say that a rock possesses some form of experientiality is not to imply that it will have a rich mental life. It won't, for instance, have a sense of selfhood, possess memories, or have the capacity to think and reason as intelligent creatures do. Rather, panpsychists merely affirm that everything which is physical possesses experiential properties. Which, when constitutive of a plant, dog, or a human, realise *what it is like* to be that particular substance.[29]

5 Anne Conway: The Panpsychist Interpretation

5.1 Matter as Intrinsically Mental

Conway's monistic substance is jointly constituted by physical and experiential properties. She writes in Chapter VI of the *Principles* that: '. . . indeed every Body is a Spirit, and nothing else, neither differs any thing from a Spirit [. . .] so that this distinction is only modal and gradual, not essential or substantial.'[30] In this sense, material substances are numerically identical to mental substances; they are one and the same thing.

Like Conway, panpsychists advocate a form of monism: viz. materialism (or physicalism). On this account, matter is posited as the basic constituent of the natural world. Furthermore, panpsychists make no essential distinction between mental and material substance. Consequently, all substances possess both experiential and physical properties. On their analysis, matter is intrinsically experiential; so that, by virtue of the nature of matter, *all* substances are constituted by basic experiential properties.

Strawson, a contemporary advocate of panpsychism, notes that one is led to the view through accepting three propositions: (a) matter is a phenomenon which exists in the natural world; (b) consciousness is a phenomenon which exists in the natural world; and (c) there is only one *type* of thing in the world.[31] If one accepts these claims,

---

27. David Chalmers, "Panpsychism and Protopsychism," in *Panpsychism: Contemporary Perspectives*, ed. Bruntrup et al. (Oxford University Press, 2016), 19.
28. Chalmers, *The Conscious Mind*, 295.
29. *Ibid*.
30. Conway, *The Principles of the Most Ancient and Modern Philosophy*, 190.
31. Galen Strawson, 'Realistic Monism', in *Consciousness and its Place in Nature*, ed. Freeman (Exeter:

then one is naturally led to the conclusion that whatever constitutes substance must be both physical and mental (in the sense of instantiating experiential properties).

In Chapter XI of *The Principles*, Conway writes that the Cartesian and Hobbesian analysis of matter in purely physical terms: '. . . profits nothing [. . .] for they have never proceeded beyond the Husk of the Shell, not reached the Kernel.'[32] In other words, they have, in their respective ways, failed to notice that since the world is monistic, and since both matter and consciousness are real phenomena, whatever constitutes the world must be both physical and experiential. Moreover, Conway asserts in Chapter VII that: '. . . every Body is a certain Spirit or Life in its own Nature, and that the same is a certain Spirit in its own nature [. . .] having Knowledge, Sense, Love, Desire, Joy, and Grief.'[33] Hence, like the panpsychists, Conway views the mental and material as constitutive of a monistic substance, and accounts for the former through the intrinsically experiential nature of the latter.

5.2   Hierarchy of Mentality

For Conway, the hierarchy of mentality is determined by the degree to which a substance instantiates Spirit (which I have argued should be understood as coextensive with 'consciousness'). Admittedly, Conway's story has more theological implications than most contemporary panpsychists would admit. Nevertheless, the moral remains: the basic constituents of the world possess some degree of experientiality, with a greater degree of experientiality being instantiated the further up the hierarchy one inspects. In fine, the more Spirit instantiated, the greater mental complexity realised.

Chalmers' discussion of panpsychism seems to suggest a similar continuum of consciousness throughout the natural world.[34] He claims that experiential properties are instantiated by all entities, with the combination of these properties realising their most complex instantiation in human minds. In other words, degrees of phenomenal consciousness are realised even by the most unthinking substances (e.g. thermostats, rocks, tables, etc.). And, by virtue of their experiential properties, there is *something it is like* to be those entities.

6   Conclusion

In short, I have shown that Anne Conway can be interpreted in such a way that she is implicitly arguing for panpsychism: the view that all substances in the world are

---

Imprint Academic, 2006), 7.
  32. Conway, *The Principles of the Most Ancient and Modern Philosophy*, 225.
  33. Conway, *The Principles of the Most Ancient and Modern Philosophy*, 191.
  34. Chalmers, *The Conscious Mind*, 293-7.

intrinsically mental. This was illustrated by showing that Conway's concept of Spirit denotes the same experiential properties as our contemporary concept of consciousness, and by tracing the similarities between her analysis of substance alongside contemporary panpsychists' analysis of matter.

This interpretation, I hope, will allow for a dual-appreciation of Conway *qua* subject of scholarly investigation, and Conway *qua* panpsychist with noteworthy ideas pertaining to the metaphysics of mind.

References

Block, Ned. "On a Confusion About a Function of Consciousness." *Behavioural and Brain Sciences* 18 (1995): 227-247.

Chalmers, David. "Panpsychism and Protopsychism." In *Panpsychism: Contemporary Perspectives,* edited by Bruntrup et al, 19-45. Oxford University Press, 2016.

Chalmers, David. *The Conscious Mind*. Oxford University Press, 1996.

Conway, Anne. *The Principles of the Most Ancient and Modern Philosophy*. Introduced by Lopston. London: Martinus Nijhoff, 1982.

Jackson, Frank. "Epiphenomenal Qualia." *Philosophical Quarterly* 32 (1982): 127-136.

Leibniz, Gottfried Wilhelm. *Philosophischen Schriften.* Edited by Gerhardt. Hildesheim: Olms, 1960.

Merchant, Carolyn. "The Vitalism of Anne Conway: Its Impact on Leibniz' Concept of the Monad." *Journal of the History of Philosophy* 17 (1979): 255-269.

Nagel, Thomas. "Panpsychism." In *Mortal Questions*, 181-195. Cambridge University Press, 1979a.

Nagel, Thomas. "What is it like to be a bat?" In *Mortal Questions*, 165-180. Cambridge University Press, 1979b.

Seager, William. "Panpsychism." In *The Oxford Handbook of Philosophy of Mind*, edited by McLaughlin et al, 206-222. Oxford University Press, 2009.

Strawson, Galen. "Realistic Monism." In *Consciousness and its Place in Nature*, edited by Freeman. Exeter: Imprint Academic, 2006.

Tye, Michael. "Philosophical Problems of Consciousness." In *Blackwell Companion to Consciousness,* edited by Velmans et al, 17-31. Oxford: Blackwell, 2006.

The Interest Theory of Rights and Rights Attached to Roles

Karri Heikkinen*
University of Glasgow

Abstract   In this essay, I defend the interest theory of rights against the argument that because some agents hold rights grounded in a special role or office, rather than their own interests, the theory is false. I begin by explaining the interest theory and the argument from rights attached to roles against it. I then explore responses to this objection by Kramer and Raz, ultimately rejecting them and proposing a simpler solution of my own. My response stipulates that in cases where people seem to have rights grounded in a special role, it is actually the state or the society that holds those rights. Given that the interest theory can allow collective entities to have rights, this response manages to retain the full explanatory power of the original theory, without introducing additional theoretical conditions for agents holding rights. Based on this, I argue that the argument from rights attached to roles is not sufficient to refute the interest theory.

1   Introduction

The aim of this essay is to defend the interest theory of rights against the argument that, because some agents hold rights grounded in a special role or office, the theory is false. The essay has three sections. In the first section, I explain the interest theory and the argument from rights attached to roles against it. In the second section, I then explore, and ultimately reject, responses to this objection by prominent interest theorists Kramer and Raz. This is because Kramer's response limits the explanatory power of the interest theory too much, whereas Raz's solution contradicts the original interest theory and necessitates the addition of new problematic conditions for agents

---

holding rights. In the final section, I propose a simpler solution of my own. This solution stipulates that in cases where people seem to have rights grounded in a special role, it is actually the state or the society that holds those rights. I argue that this response allows us to explain many cases of rights related to special roles that are often thought to be difficult for the interest theory to accommodate. Given that the interest theory can allow collective entities to have rights, this response manages to retain the full explanatory power of the original theory, without introducing additional theoretical layers. Finally, I argue that the idea that collective entities can have rights is not merely a peculiar feature of the interest theory, but present in many areas of our common language. Based on this, I conclude that the argument from rights attached to roles is not sufficient to refute the interest theory.

2   The Interest Theory and the Argument from Rights Attached to Roles

The interest theory is a theory about the normative foundation of rights. For the purposes of this essay, I am trying to avoid as many additional commitments related to the definition of rights as possible. Rather, I take "rights" as a basic moral concept which can be explicated in various ways by different first-order moral and political theories, similarly to terms like ''duty'' or "good". Although different authors have presented slightly different formulations of the theory, a simplified version of Kramer's interest theory will be sufficient for the purposes of this essay. Paraphrasing Kramer[1] [2], the interest theory consists of two main propositions:

1. For X to hold a right, it is necessary, though not sufficient, that the right protects some aspect of X's situation that is typically beneficial for a being like X.

2. For X to hold a right, it is neither necessary, nor sufficient, that X is competent to demand or waive the enforcement of the right.

Proposition 1 states the basic idea of the interest theory, namely that rights exist to make right-holders better off. For example, because it is very much in our interest to not be physically assaulted, we have a right to bodily integrity. The point of proposition 2, on the other hand, is to make the claim that having a right does not depend on the right-holder having highly developed mental competencies. What is required is simply that the right-holder has interests, not that they are autonomous, for example. Therefore, the interest theory entails that the group of potential right-holders is not limited to

---

1. Kramer, M.H., "Some Doubts about Alternatives to the Interest Theory of Rights",'*Ethics*,vol. 123, no. 2, (2013): 246.

2. Kramer, M. H., Simmonds, N. E. & Steiner, H., *A debate over rights: philosophical enquiries* (Oxford: Oxford University Press, 1998): 62.

competent adult humans, but can include any entities that have interests, for example non-human animals or corporations.

It is important to note that the interest theory does not claim that all interests create rights: to give an obvious example, even if an aspiring politician has a strong interest in being elected, it does not follow that they should have a right to be elected regardless of their expertise and support among their constituents. This is why having an interest is a necessary, but not a sufficient condition for having a right. Indeed, the interest theory, as defined above, does not take a stance on what set of conditions would be *sufficient* for having a right, allowing the theory to be paired with a variety of moral views. For example, it seems plausible to say that for an interest to ground a right, the interest must not involve gratuitously harming others, meaning that the interest theorist need not be committed to claiming that, say, someone who gets pleasure from killing dogs must have a right to do so. However, discussing exactly what set of interests and possible other conditions would be sufficient for creating a right is beyond the scope of this essay. Instead, my aim is to defend a far more limited, but still an important claim about the source of rights, namely that for every right someone holds, there is a corresponding interest that the right protects. As we shall see, some philosophers think that there are obvious counterexamples against this claim, which I aim to refute.

Before moving on, however, it is useful for understanding the interest theory to briefly think about the positive reasons there are for endorsing it. To begin with, it seems very intuitive to think that many rights exist because they make us better off. In addition to the right to bodily integrity, rights such as right to life, right to subsistence and right to education seem so important namely because they are vital to our welfare. A related point is that the interest theory conceptually allows there to be unwaivable rights. For example, interest theorists can maintain that people have the unwaivable right to not be enslaved because this is always in their interest. On the other hand, the interest theory is not committed to the claim that our rights always make us better off in every instance: for example, as Wenar[3] points out, inheritances might sometimes cause us more trouble than benefit. This is why proposition 1 in our definition of the interest theory talks about things that are *typically* beneficial for beings like the right-holder in question. Finally, it seems very plausible that beings other than adult humans, such as infant human children, have at least some important rights. The interest theory has no difficulty in accommodating this intuition.

Having introduced the interest theory, I will now present the argument from rights attached to roles against it. The idea is simple: contrary to what the interest theory claims, there seems to be cases where people possess certain rights due to a special role or office that they hold. An oft-cited example is a judge who has the right to convict people to years in jail.[4] I will reconstruct the argument in standard form as follows:

---

3. Wenar, L., "The Nature of Rights", *Philosophy* & *Public Affairs*, vol. 33, no. 3 (2005): 241.
4. *Ibid.*, 242.

P1 The interest theory maintains that for X to hold a right, it is necessary that the right protects some aspect of X's situation that is typically beneficial for a being like X.

P2 The right of a judge to convict people does not protect any aspect of the judge's situation that is typically beneficial for humans (i.e. beings like the judge).

C1 It is not necessary for X to hold a right that the right protects some aspect of X's situation that is typically beneficial for a being like X.

C2 The interest theory is false.

The argument uses a counterexample to show that proposition 1 from our definition of the interest theory, as expressed in **P1** in the argument, is false. Because the right of the judge to convict criminals does not seem to be relevant for any important aspect of their welfare or other interests - unlike their right to bodily integrity typically is - it seems impossible that the interests of the judge could give them the right to alter the situations of other people so dramatically. Instead, the natural interpretation of the case is that the judge is exercising a right that is not based on their interests. Therefore, it seems that having an interest is not a necessary condition for holding a right.

3 Responding to the Argument

The argument from rights attached to roles has been discussed extensively in the literature. Since my aim in this essay is to defend the interest theory, I will now look at some responses that have been offered by prominent interest theorists to see if they can aid, or at least inform my project. To begin with, Kramer[5] responds to the argument from rights attached to roles by maintaining that, in fact, the right of the judge to convict criminals is not a right at all. According to Kramer, when we talk about the right of a judge to convict or a head of state to declare war, we are talking about powers or special capabilities that are simply not the same thing as rights. I find this solution somewhat intuitive, as I feel that although the same word is commonly used, the right of a judge to convict criminals is a very different case from ordinary rights that belong to everyone, such as the right to basic sustenance. Perhaps we should reserve the concept of right to specific usage and find more suitable terms for the "rights" of office-bearers.

Still, I think that accepting this response would strip away an important part of the interest theory's explanatory power. This is exactly because Kramer's response drastically limits the types of rights that the interest theory can accommodate, accepting only claims as actual rights and leaving out all other Hohfeldian incidents, namely

---

5. Kramer, M.H., Simmonds, N.E. & Steiner, H., *A debate over rights: philosophical enquiries*, 9-10.

privileges, powers and immunities.[6] I find that a better way to account for the sense of important difference between the judge and someone's right to bodily integrity is to say that the judge holds a power-right to convict, whereas we all hold claim-rights against bodily attacks. The difference here is, roughly, that the judge has a right to alter other people's (claim-)rights (in this case, the convicted person's right to liberty), whereas our right to bodily integrity just means that others have a duty not to attack us. While it is not the purpose of this essay to dwell deep into Hohfeld's terminology, it should nevertheless be clear that not being able to accommodate many of the common-sense ways we refer to rights in our language is a flaw for a theory seeking to explain them. Of course, this flaw could be rectified with a convincing positive argument for why we should understand rights exactly in the way Kramer suggests, but with no such argument in sight, I think it is best to keep refraining from additional normative assumptions regarding the exact nature of rights.

At first glance, I find that Raz[7] offers a more promising solution. He holds that rights are not always fundamentally grounded on the personal interest of the right-holder. Sometimes an interest can be instrumentally valuable because it promotes the welfare of others. For example, while it is true that the right of a judge to convict criminals cannot be grounded on the judge's personal interests, this right can be justified by appealing to the interests of the population in general. Here, the judge has an interest in being able to do their work, but this interest is sufficient to create rights only because it promotes safety, and as such is instrumentally valuable to other moral agents living in the society. Raz' own example is a journalist, whose right to protect the identity of their sources is ultimately justified because of its value to the public.[8]

However, Kamm[9] has pointed out that Raz' argument contradicts the interest theory as it is commonly defined. From our point of view, it is fairly clear why this is: if it is not the interests of the judge, but those of someone else, that give the judge their power-rights, then proposition 1 in our definition of the interest theory is false. This is because it is no longer necessary for X to have a right that this right protects some aspect of namely X's own situation. To make the interest theory compatible with Raz' solution, our definition would have to be modified like this:

1.* For X to hold a right, it is necessary, though not sufficient, that the right protects some aspect of X's situation that is either **a)** typically beneficial for a being like X or **b)** benefits a large enough group of other morally significant beings.

It is important to unpack this definition carefully. What is proposed here is that while it is still X that holds the right, this right need not be grounded in X's interests.

---

6. Hohfeld, W.N., "Fundamental Legal Conceptions as Applied in Judicial Reasoning", *Yale Law Journal*, vol. 23, no. 1 (1913): 28-58.

7. Raz, J., The morality of freedom (Oxford: Oxford University Press, 1988): 178-190.

8. *Ibid.*, 179.

9. Kamm , F., "Rights" in *The Oxford handbook of jurisprudence and philosophy of law*, ed. Coleman, J.L. & Shapiro, S. (Oxford: Oxford University Press, 2004): 485.

Instead, X can hold a right that protects the interests of [Y, Z... N]. For example, the journalist's right to protect the identity of their sources in Raz' example is still a right that the journalist holds personally, but the moral grounding of this right is in the interests of other people. This represents a major change compared to the original interest theory, where agents can only hold rights grounded on interests that they hold themselves.

The fact that this modification changes the interest theory substantially leads to some issues. Firstly, it seems that the added part b) of the proposition effectively allows consequentialist considerations to create rights. If one believes that consequentialist calculations are an inappropriate or even impermissible way of arriving at moral judgements, then this might be a fatal flaw in itself. However, even if we allow our moral theory to assign some value to consequences, this formulation of the interest theory clashes with many people's intuitions of rights, as rights are usually considered in some sense unalterable, meaning that they cannot be erased whenever an individual's rights clash with the interests of the majority. Secondly, it is also not clear which of the two functions that rights can serve in the revised interest theory should take priority if they contradict each other. For example, it can definitely be in a rich person's interests to have strong property rights, whereas the public benefit might require that the state has the right to tax its citizens. These kinds of evaluative problems are of course present in the original interest theory but adding multiple functions that rights can fulfil makes them even more difficult. Finally, even if we managed to solve these problems, this solution adds theoretical layers into the interest theory, making it less parsimonious. Therefore, it seems that this solution is, after all, only a minor improvement to Kramer's response.

4   Collective Entities as Right-Bearers

Despite these problems, I find that there is something correct about the idea that the public interest can ground rights. Therefore, I will now try to develop this idea. To tackle the objection from rights related to office, while avoiding the problems described in the preceding section, I think a supporter of the interest theory should go back to the fundamental properties of their theory. Here, I believe that the interest theorist can make use of the fact that the interest theory does not limit the group of potential right-holders to adult humans, but can also include entities like corporations or states. In cases like the judge example, the interest theorist could simply say that it is the state that holds the right to convict criminals, rather than the judge. The judge is merely a tool that follows carefully defined rules, allowing the state to exercise its right to punish people who harm it. This simple solution can still accommodate all Hohfeldian incidents and, unlike Raz' idea, it does not require that we change proposition 1 of our definition of the interest theory. This is because unlike in Raz' solution, we do not

claim that the judge holds any power-rights at all, meaning that we need not justify the rights of one agent by the interests of someone else. Instead, both the right-holder and the interest-holder, in this particular case, is the state. It is not the case that X can hold rights grounded in the interests of [Y, Z. . . N] but, instead, X simply *is* [Y, Z. . . N].

To see whether my suggestion seems plausible, we can try to apply it to some other cases. To begin with, I think we can say something along my line of thought about the journalist protecting their sources. It seems plausible to me that, in fact, the journalist does not have a right to protect their sources, but a duty to do that. Consequently, we can analyse the case as an instance of the civil society as a whole having a claim-right that creates this duty. If the journalist would harm the civil society by naming their sources, then it is conceivable that the civil society has a right against this kind of harming. In this vein, it seems like most arguments that at first seem to make references to multiple other people's interests as a way to justify someone's rights could be phrased in terms of collective entities, by simply identifying the collective that the interest-holders form. Nevertheless, there still remains some difficult cases where it seems like someone holding a special role has the right to act against the interests of the relevant collective. For example, consider a head of state who has a constitutional right to order a nuclear attack, and consequently could do this because of a personal whim, even if it was against the national interest. However, I think it is clear that even if this head of state would have the legal right to order the attack, they would not have the moral right to do so. In fact, in this case, I hold that the nation would have a moral right against the head of state ordering nuclear attack. There is nothing peculiar here: since the interest theory is a theory of moral rights, there will be cases where the legal rights of certain office-bearers clash with the moral rights of both individuals and collectives. Making this distinction allows us to explain why it can sometimes seem as if certain office-bearers have the "right" to act in ways that are against everyone else's interests, even though this is not the case in so far as we are concerned with morality rather than the letter of the law.

Having established that my proposal can make sense of many situations that are usually considered difficult for the interest theory to accommodate, I wish to address one more possible objection to my view. I believe one could make various metaphysical arguments aimed at establishing that the kind of collective entities that I refer to do not really exist, and therefore, my view is not sound. As a pre-emptive, though imperfect, rejoinder to this class of objections, I would like to point out that we already seem to accept, in many areas of life and our common language, that many kinds of collective entities can hold rights. A paradigm example of this is a corporation with copyright to intellectual property. Moreover, in international law, there are references to rights of states,[10] and in the age of climate change, more and more references are made to rights

---

10. See e.g. Brown, P.M., "The Rights of States under International Law", Yale Law Journal, vol. 26, no. 2 (1916).

of future generations.[11] Of course, this does not *prove* that collective rights exist, in the same way that the fact that we can talk about dragons does not mean that dragons exist. However, given the intelligibility and common occurrence of the idea that collective entities can have rights, I hold that even my dialectic opponent should grant that this idea seems initially plausible. Therefore, the burden would be on them to explain what goes wrong when we talk about collective rights and why these rights do not really exist. It is also true that deciding exactly what kind of rights, if any, a certain group should have and how these rights should weigh against rights of others is a difficult question, but this does not undermine the general idea that collective entities can hold rights.

5 Conclusion

In this essay, I have defended the interest theory from the argument from rights attached to roles. I began by introducing the interest theory and the argument against it, before reviewing two responses by prominent interest theorists Kramer and Raz. Instead of endorsing either of them, however, I argued that we can better respond to the argument by stating that it is actually the state or the society that holds the rights we often erroneously attribute to office-bearers. This response has the benefit of retaining the interest theory's explanatory power across all Hohfeldian incidents and avoiding a collapse to utilitarianism, without having to introduce additional theoretical conditions in it. After all, it is a central feature of the interest theory that the group of potential right-holders is not limited to individuals, but can include collective entities. We also saw that my solution can successfully explain many cases where agents seem to have rights grounded in holding certain roles, as well as allowing us to say what is wrong with cases where office-bearers seem to have a right to act against everyone else's interests. Finally, I pointed out that the idea that collective entities can hold rights is not confined to the literature regarding the interest theory, but instead, can be found in our common language from many areas of life. This observation lends further support for the plausibility of my view against those who are sceptical of the existence of the collective entities I refer to. Therefore, I conclude that the argument from rights attached to roles is not sufficient to refute the interest theory.

References

Brown, Philip Marshall. "The Rights of States Under International Law" *The Yale Law Journal* 26, no. 2 (1916): 85-93.

---

11. For some suggestions, see Gosseries, A., "On Future Generations' Future Rights", *Journal of Political Philosophy*, vol. 16, no. 4 (2008).

Gosseries, Axel. "On Future Generations' Future Rights". *Journal of Political Philosophy* 16, no. 4 (2008): 446-474.

Hohfeld, Wesley Newcomb. "Fundamental Legal Conceptions as Applied in Judicial Reasoning". *The Yale Law Journal* 26, no. 8 (1917): 710-770.

Kamm, Frances. "Rights". In *The Oxford handbook of jurisprudence and philosophy of law,* edited by Jules L. Coleman, Kenneth Einar Himma and Scott J. Shapiro, 476-514. Oxford: Oxford University Press, 2004.

Kramer, Matthew Henry. "Some Doubts about Alternatives to the Interest Theory of Rights". *Ethics* 123, no. 2 (2013): 245-263.

Kramer, Matthew Henry, Nigel E. Simmonds, and Hillel Steiner. *A Debate Over Rights: Philosophical Enquiries*. Oxford: Oxford University Press, 1998.

Raz, Joseph. *The Morality of Freedom*. Oxford: Oxford University Press, 1988.

Wenar, Leif. "The Nature of Rights." *Philosophy & Public Affairs* 33, no. 3 (2005): 223-252.

A Metacognitive Approach to Memory Markers

Matheus Diesel Werberich*
Federal University of Santa Maria, Brazil

Abstract   Given both the phenomenological and cognitive similarities between episodic memory and imagination, it's difficult to say how we can reliably distinguish them at their moment of retrieval. Several memory markers have thus been proposed, which are characteristics that would reliably indicate to the subject that her mental state is an instance of memory. While the question of what exactly constitutes these memory markers is still an issue to be settled, there is also the more general question of whether they can be reliable at all. In the present paper, I have identified two theses about the latter issue (the reliability and unreliability theses) and have argued that the main cause of disagreement between them lies on their different assumptions on how beliefs about our own mental states are formed, i.e. what are the underlying metacognitive mechanisms responsible for self-attribution of mental states. These different views on metacognition were then further investigated with regards to their use of metarepresentations and their general agreement with recent cognitive psychology research. These analyses corroborate the reliability of certain kinds of memory markers in distinguishing between memory and imagination.[1]

1   Introduction

There are many different metaphors used to explain memory in the history of western philosophy, from Plato's wax tablet to Locke's storehouse imagery. What all of these conceptions have in common is that they take memory to be a passive system:

---

information is carried through perception to memory, where it is stored for later re-
trieval. Although this characterization was used for a few centuries, this framework
is no longer widely adopted to account for episodic memory, i.e. the recollection of
personal past episodes. This is mainly due to recent findings in neuroscience and cog-
nitive psychology.[2] Researchers in those areas discovered that a major part of episodic
memory's neural substrates are the same as those of episodic imagination: the same
areas and processes are active when the subject is remembering a past episode and
when she is imagining a future or counterfactual event. Thus, it has been widely ac-
cepted that episodic memory, like imagination, is a kind of process that involves the
active construction of episodes, not the passive retrieval of them from past experience.

Given its constructive character and relation to episodic imagination, many philo-
sophers have claimed that episodic memory and imagination share the same nature, so
much so that there is not a qualitative distinction between them. This proposal, called
continuism, is at odds with the more traditional and commonsensical view, named dis-
continuism, in which episodic memory and imagination are different in nature, even
though their phenomenology might be similar.[3] Whichever is the true alternative, they
both face the same problem: if memory and imagination are so phenomenologically
similar, then how can we tell them apart at the moment of retrieval? Is there an element
that would be exclusive to, or at least largely more present in memory, that distin-
guishes it from imagination?[4] Traditionally, the answer to the latter question has been
affirmative, and philosophers of memory call this kind of element a *memory marker*. In
what follows, I will analyze this kind of proposal and argue that the discussion about
the reliability of these memory markers depends on certain considerations about the
functioning of a larger metacognitive system.

2   The Memory Markers Proposal

In general, there are two kinds of memory markers: first person markers and third
person markers. The first kind are the ones to which the subject has direct access at
the moment of retrieval. One classic example of this kind is the feeling of familiarity:
the content of memory representation feels familiar to the subject, giving her the con-

---

2. For an example of such studies, see Schacter and Donna Addis, "The cognitive neuroscience of
constructive memory". Philosophical Transactions of the Royal Society of London. Series B, Biological
Sciences 362, 1481 (2007). doi: 10.1098/rstb.2007.2087.

3. For a more detailed discussion on continuism and discontimuism, see Kourken Michaelian and
Denis Perrin, "Memory as mental time travel" in The Routledge Handbook of the Philosophy of
Memory, ed. Sven Bernecker and Kourken Michaelian (New York: Routledge, 2017): 228 – 239.

4. Note that this is a different issue from how the agent can trust that her memory is in fact giving
her reliable information about the past. The solution to this problem relies not on memory markers, but
rather if the memory process itself can reliably retain, or reconstruct, information from the past.

fidence to consider it as a case of episodic memory. Meanwhile, third person memory markers are external to the subject's current mental state because they can only be determined by analyzing certain aspects that concern memory's formation process. For example, an unbroken causal chain between the original experience and its current representation is a kind of third person memory marker because only memory can have it, and it cannot be detected just by checking its resultant mental state. Since philosophers consider that we can tell memory from imagination at the moment of retrieval, their focus will be to search for first person memory markers, which, following Michaelian's classification, can be divided into three groups: formal, content-based, and phenomenological.[5]

Formal memory markers are concerned primarily with the way that memory content is presented to the agent. For instance, memory tends to be more spontaneously retrieved than imagined scenarios, so much so that the speed of memory's retrieval can be an indicator that the subject encountered something similar in the past. Other formal criteria include memory's inflexibility (i.e. memory's susceptibility to change based on other inputs, such as other's testimony), its spontaneity, fluency (i.e. the relative easiness of memory's formation, related to the spontaneity criterion) and the intention to remember.

In turn, content-based memory markers pertain to the idea that certain aspects of memory content can be used as criteria for its classification. Characteristics such as greater level of detail and coherence with other beliefs may play a role in distinguishing between memory and imagination. It is important to note, as it will become relevant in future sections, that these two categories of memory markers, i.e. formal and content-based, aren't a matter of absolute presence or absence, but rather are characteristics that are more present in memory than in imagination.

The third and final type of memory markers is concerned with the phenomenological aspects of episodic memory. In particular, these accounts consider the metacognitive feelings brought by memory that should be absent in cases of imagination. Metacognitive feelings are a way of the metacognitive system to point certain features about other cognitive processes. Such feelings include the feeling of prior belief, pastness, familiarity, and confidence. These criteria are not exactly characteristics of the memory process, but are rather generated by the metacognitive system through its monitoring of the episodic construction process.

---

5. Kourken Michaelian, Mental Time Travel: episodic memory and our knowledge of the personal past. (Cambridge, MA: The MIT Press, 2016), 181 - 194.

3   The Metacognitive Approach to Memory Markers

Now that we have seen the major proposals for memory markers found in the literature, we can turn our attention to a more general problem: can memory markers generally be reliable in distinguishing between memory and imagination? The answer to this question will largely depend on what counts as "reliable" : if the motivation provided by the memory markers comes from a reliable enough system, then the answer would be affirmative; on the other hand, if we define reliability as the capacity to produce justified beliefs based on well-grounded criteria, then memory markers wouldn't be the most appropriate, since most of them can be also found in imagination. Therefore, we have two theses about the overall reliability of memory markers:

(1) Reliability thesis: the motivation to classify a mental state as memory provided by memory markers is sufficiently reliable, since it comes from an overall reliable system;

(2) Unreliability thesis: the motivation provided by memory markers is not sufficiently reliable, since they don't provide valid justification for metamemory beliefs.

In the recent literature, these two theses are defended mainly by philosophers standing in different camps in philosophy of memory. Simulationists, such as Michaelian[6], would generally argue for the reliability thesis, since, if episodic memory is defined as reliable imagining, or simulation, of past events, its reliability would be lost if the agent couldn't distinguish between it and other forms of imagination. According to Michaelian, "one normally 'just knows' whether one is remembering or imagining"[7], thus pointing to the fact that an active and conscious introspection might not be required for determining whether an episodic state is a case of memory or imagination.

On the other hand, causal theorists, like Bernecker[8], tend to argue for the unreliability thesis, stating that the motivation provided by memory markers to distinguish between memory and imagination does not mean that they provide "an epistemic reason for anything" [9]. This means that the result of using memory markers should be a justified belief about one's episodic mental state, not just some notion or feeling, as in the reliability thesis.

One thing to note about the two theses is that their considerations on the reliability of memory markers come from different views on how metamemory beliefs are formed. The reliability thesis considers that there is a subconscious background system which closely monitors the memory process and, depending on the characteristics detected, makes the subject more susceptible to treat her mental state as memory. On the other hand, the unreliability thesis takes the metamemory beliefs as forming through

---

6. Kourken Michaelian, Mental Time Travel: episodic memory and our knowledge of the personal past.

7. *Ibid*, 194.

8. Sven Bernecker, Memory: a philosophical study (Oxford, Oxford University Press, 2010).

9. *Ibid*, 33.

a sort of introspection, where the subject will consciously look for memory markers in her mental states and form a judgement based on their presence or absence. Thus, the unreliability thesis requires that the subject has at least some idea of how to classify her mental states, while the reliability thesis does not require such mental capacity, only that a background subsystem must be able to detect certain features of memory.

Therefore, the discussion on the general reliability of memory markers mainly comes from different views about how mental states can be detected and evaluated. These capacities are studied in the field of metacognition, which is still characterized by a general lack of agreement in many issues. In order to organize these problems, Joëlle Proust[10] distinguishes between two general theories of metacognition: the evaluativist view and attributive view. According to an attributive view of metacognition, the subject becomes aware of her own mental states through a conceptual representation of them. Thus, the metacognitive system would generally rely on what is similar to type 2 kind of processing[11], which is regarded as the slow, conceptual and propositional mechanism of information processing. In contrast, the evaluativist view, also known as procedural view, considers metacognition to largely rely on what is similar to type 1 processing, which is the fast, subconscious and heuristics-based mechanism, responsible for the evaluation of general features of cognitive processing, such as fluency.

If one considers the evaluativist view to be true and its monitoring system to be reliable enough, then one naturally comes to the conclusion that the reliability thesis is true. Conversely, if one accepts the attributive view, then the memory markers presented in section 2 may not serve as a sufficient justification factor for their metamemory beliefs, thus following the unreliability thesis. In what follows, I will present in more detail how the evaluativist and attributive views work, as well as their implications for the question about the general reliability of memory markers.

3.1 Attributive Views

In the attributive view, any form of metacognition requires a metarepresentation of the target mental state. Metarepresentations are understood as a representation of both the content and the representational vehicle. For example, if a subject has a mental picture of her future birthday party, then her metacognitive system would have to represent the same mental picture *as* a case of imagination – her metarepresentation would have information of both the content and the process involved in making it[12]. Following

---

10. Joëlle Proust, The Philosophy of Metacognition: Mental Agency and Self-awareness (Oxford: Oxford University Press, 2013).

11. For a more detailed account of these kinds of cognitive processing, see Daniel Kahneman, Thinking, Fast and Slow (New York: Farrar, Straus and Giroux, 2011).

12. For a more detailed discussion on metarepresentations, see Josef Perner, "MiniMeta: in search of minimal criteria for metacognition", Foundations of metacognition, ed. Michael Beran, Johannes Brandl

this, many researchers conclude that metacognition employs the same mechanisms of mindreading, which is the capacity to attribute mental states to other people based on perceptual clues, such as facial expressions, tone of voice, bodily movements, among many others. If that is the case, then attributive theorists would have to explain what exactly the relation between mindreading and metacognition is, which would in turn have consequences for their general theory of introspection.

For example, Carruthers claims that metacognition, or introspection, is just self-directed mindreading.[13] His claim is that perceptual and recognition mechanisms, alongside concepts of mental states and the self, would be able to extract clues from the subject's internal states and infer propositional attitudes, just as it would to other people. Such clues include inner-speech, visual imagery, attention shifts, and so on. Of course, some of these clues might not be available for attributing mental states to other people. For instance, mental imagery of other agents is not as easily detected as one's own. However, this does not pose a problem for Carruthers' account, for it still might be the case that the same mechanisms are involved both in mindreading and introspection, just not necessarily the same inputs would be used in evaluating other's or one's own mental states.

One consequence of Carruthers' model is that we only have interpretative access to some of our mental states. If the mechanisms of mindreading are based on the interpretation of clues in another agent, then, when they turn on their own subject, they would have to interpret her behavioral and perceptual clues as well. Thus, much of our self-ascription of mental states has to be interpretative, including metamemory beliefs. In this case, the metarepresentational system would receive an episodic construction as input and ascribe the concept of "remembering" to the subject.

Alternatives to Carruthers' account of metacognition and mindreading generally use the same sort of metarepresentational mechanisms, but put the metacognitive capacity as being more fundamental than mindreading. An example of this is Goldman's account of self-knowledge, in which we only acquire the concepts necessary for metarepresentations through the detection of our own mental states.[14] Based on this, the subject can then perceive behavioral clues in other agents and apply those concepts to them (i.e. the concepts of remembering, reasoning, and so on).

If metamemory beliefs are based on interpretation of clues, then memory markers can be detected for the self-ascription of episodic memory. Considering the formal and content-based markers presented in section 2, the perception of them could be used as input for the metarepresentational system. However, if memory and imagination share a lot of their formal and content-based properties, then it is unlikely that any

et al. (Oxford: Oxford University Press, 2012): 94 - 116.

  13. Peter Carruthers, "How we know our own minds: the relationship between mindreading and metacognition", *Behavioral and Brain Sciences* 32, 2 (2009): 121-182. doi:10.1017/S0140525X09000545.

  14. Alvin Goldman, *Simulating Minds: the philosophy, psychology and neuroscience of mindreading* (Oxford: Oxford University Press, 2006): 246.

single criterion, or set of criteria, can reliably set them apart – any of them would also be present in imagination, undermining the very notion of memory markers, which are characteristics that would characterize only memory states.

One might argue that this is only a problem if we adopt continuism about episodic memory and imagination, which is the view that they only differ in terms of degree, but not of kind. If one adopts discontinuism, on the other hand, they would hold that memory and imagination are different kinds of mental states, thus not being able to share formal and content-based properties in the same way that was described earlier. However, even if they are different kinds of mental states, it does not mean necessarily that they don't have common formal and content-based properties, since they are both mental images of non-present events. Despite the possibility that they are formed by different systems, the fact still remains that they present themselves in a similar way, for they are mental representations of non-present scenarios.

Therefore, notwithstanding the actual differences between memory and imagination, if one assumes an attributive view of metacognition, they would tend to consider formal and content-based memory markers as unreliable criteria for distinguishing memory and imagination. However, what if the metarepresentational system took the phenomenological markers as input? Can feelings of remembering serve as reliable criteria for episodic recollection? The answer to these questions depends on how these epistemic feelings are formed, which is the topic of the following section.

3.2   Evaluativist Views

As was stated earlier, the evaluativist view considers metacognition to be a subconscious, fast process of monitoring and control of cognitive mental states. This largely coincides with what is called type 1 processing, which is a kind of cognitive capacity that aims at achieving fast production of outputs from initial inputs. Since this is not a conscious or deliberate act of the subject, the outputs provided by the evaluativist model of metacognition will generally lack the conceptual and propositional structure that other cognitive states have.

In the evaluativist view, it is useful to distinguish between two levels of processing: the cognitive level, where mental states are primarily constituted of what ordinarily is thought of our mental activities, such as memory, beliefs, imagination and so on, and, the meta level, which is primarily concerned with tracking what is happening at the cognitive level. This kind of interaction between the two levels can happen in two different ways: monitoring, where information flows from the cognitive level to the meta level, and control, which has the opposite flow of information, allowing for the meta level to intervene on the cognitive level.

Since we are concerned with the overall reliability of distinguishing between memory and imagination, our main concern should be how exactly monitoring oc-

curs: how can the metacognitive level acquire relevant information from the cognitive processes to properly categorize them? Since categorization of mental states usually occurs through metacognitive feelings, evaluativist theories mainly focus on how the metacognitive system can monitor the cognitive level and create those subjective experiences, such as the feeling of pastness and of prior belief.

In this framework, we have two questions about memory markers. The first one is concerned with which are the memory markers that the metacognitive system would be able to detect in order to reliably classify a mental state as memory. We can call it *the monitoring question.* The second one is more concerned with later stages of the metacognitive process, asking which sort of epistemic feelings would the system generate to motivate the subject to accept the mental state as memory. We can call it *the feelings question.*

In order to answer these problems, Dokic first considers how the metacognitive system can produce epistemic feelings while trying to explain memory's phenomenology.[15] According to Dokic, episodic memory has a distinct phenomenological feature, namely the episodic feeling of knowing, through which the subject becomes more inclined to think that her mental state comes directly from the past experience. Based on empirical research, Dokic claims that the episodic feeling of knowing is based on the detection of certain features of the episodic construction process, such as fluency and familiarity, which are reliable indicators of episodic memory.[16]

Following this model, we can have some answers to the two questions about memory markers described earlier. The *feelings question* is about which markers are used in order to inform the subject about her mental states. Dokic answers this with the episodic feeling of knowing, since this feeling is a product of heuristics that are indications of the construction process being connected to the original experience. Following the classification presented in section 2, the episodic feeling of knowing would be a phenomenological memory marker.

The *monitoring question* is concerned with which are the memory markers are detected by metacognition that allow the episodic representation to be characterized as memory. Since phenomenological markers are only present in later stages of this kind of metacognitive monitoring, the decision must be between formal and content-based markers. Fortunately, empirical research has been fruitful in this regard, showing that monitoring of content of the object level does not seem to be as useful as tracking their general features.[17] Thus, formal memory markers are the best examples for what exactly the metacognitive system is detecting in order to form the episodic feeling

15. Jérôme Dokic. "Feeling the Past: A Two-Tiered Account of Episodic Memory." Review of Philosophy and Psychology 5, no. 3 (2014) doi: 10.1007/s13164-014-0183-6.

16. *Ibid*, 422.

17. Asher Koriat, "The subjective confidence in one's knowledge and judgements: some metatheoretical considerations", Foundations of metacognition, ed. Michael Beran, Johannes Brandl et al. (Oxford: Oxford University Press, 2012): 215.

of knowing. Favorable candidates include memory's spontaneity and fluency, since the easiness of memory's formation could be an indicator of the system having encountered the same contents in the past.

While Dokic's metacognitive model, as well as other evaluativists accounts, can ground the overall reliability of memory markers, they have one specific problem to be solved. According to evaluativism, the metacognitive system is relatively simple when compared with the cognitive system, since it does not use concepts or metarepresentations for monitoring and control. If that's the case, then procedural evaluativists would have to explain how the metacognitive level is able to tell, without relying on metarepresentations or concepts, that some elements of the episodic construction process indicate features of memory. This is relevant for the overall reliability of phenomenological memory markers, for their adequacy is dependent on how they are formed.

One possible way of answering this problem is by relying on what Proust calls Test-Operate-Test-Exit (TOTE) units,[18] which are feedback loops designed to compare a current mental state with the desired one (called the comparator) and, based on that comparison, inform control about what sort of action is necessary. For instance, when monitoring a given memory process, the metacognitive system would compare the episodic process' fluency with the previously stored comparator. If no major discrepancy is detected, then a feeling of remembering would be generated. This mechanism can avoid having conceptual information about the cognitive process, since the comparator does not necessarily need to have a concept of fluency clearly stated in order to operate: if the system is sensitive enough to the difference in speed and easiness between the formation of the current mental state and its comparator, then no concept of fluency is required. Following this, it can also avoid having to deal with metarepresentations: since they are defined as the representation of mental states as such, they necessarily require concepts about those cognitive states. Once the system does not need concepts in order to operate, then it can also do without metarepresentations.

Appealing to empirical research, Proust further shows that procedural metacognition does not need to entail metarepresentations. Research on non-human animals has indicated that procedural metacognition is likely to occur with some species of primates, which shows that "noetic feelings can monitor animals' decisions to act on a memory or on a percept, even though the concepts and percept aren't available to them".[19] Since nonhuman animals can reliably assess their probability in resolving certain cognitive tasks while also not being able to mindread (at least not in the same conceptual way as humans do), this is a strong indication that metacognition does not require metarepresentations or concepts in order to be performed properly.

Furthermore, studies in cognitive psychology show that there is a strong relation

---

18. Proust, The Philosophy of Metacognition: Mental Agency and Self-awareness: 14.
19. *Ibid*, 98.

between the fluency of the retrieved memory and the accuracy of the subsequent feeling of knowing. Asher Koriat, for example, states that "with regard to the accuracy of metacognitive judgements, the observation that has attracted the attention of researchers in metacognition is that participants are generally accurate in monitoring their knowledge".[20] Since these metacognitive judgements are primarily based on feelings of knowing, we can infer that the same is true of monitoring episodic memory: the cues extracted in the memory process, such as some formal memory markers, are sufficient for the formation of a subsequent feeling of remembering, which reliably prompts the agent to endorse the information as memory, not imagination. Therefore, it appears that empirical research corroborates both the procedural model of metacognition and the reliability thesis about formal and phenomenological memory markers: the former would constitute reliable heuristics to be monitored, while the latter can be seen as a good indicator to the subject about her episodic mental state. With regards to content-based markers, since they cannot be directly monitored by procedural mechanisms, neither serve as reliable input for the attributive system (as was seen in section 3.1), we can conclude that they are not reliable criteria for episodic memory.

4   Summary and Conclusion

Given that episodic memory is so phenomenologically similar to imagination, philosophers have long been debating how we can reliably distinguish them at the moment of retrieval. One of the most popular solutions is to posit certain elements that would characterize the presence of episodic recollection, called memory markers. As we saw in section 2, those can be classified into three groups: formal, content-based and phenomenological markers. Furthermore, philosophers have also debated whether memory markers can be reliable at all. The present paper divided this discussion between two opposing theses, the reliability and the unreliability theses, and showed that their main cause of disagreement is that they make different considerations about the formation of metamemory beliefs, which in turn depend on their views about metacognition in general.

The unreliability thesis considers metacognition to be an attributive and conceptual process. In this view, memory markers would serve as possible criteria for the agent to detect and form a metarepresentational belief. As we saw in section 3.1, if one uses attributive mechanisms for self-attribution of episodic memory, then formal and content-based memory markers should be rejected for they don't provide enough justification or certainty for the subject to reliably form her metamemory beliefs.

On the contrary, the reliability thesis takes metacognition to be a procedural and evaluative system, in which mental states are evaluated through a fast and subcon-

---

20. Asher Koriat, "The subjective confidence in one's knowledge and judgements: some metatheoretical considerations", 213.

scious process. With regards to memory markers, their overall reliability would depend on the reliability of the metacognitive system, which aims to detect formal markers in the memory process and to generate an epistemic feeling based on their presence. A model of the detection of these markers was presented based on the concept of "Test-Operate-Test-Exit" units. These metacognitive mechanisms would be capable of comparing the detected fluency of the memory system against a certain threshold. If the perceived value meets or exceeds the comparator, then a feeling of recollection is generated. Given that this kind of process may also be present in non-human animals, it is more likely that complex and conceptual metarepresentations aren't present in procedural metacognition. Furthermore, findings in cognitive psychology show that the results of procedural metacognition, i.e. epistemic feelings, are largely reliable in assessing the reliability of memory processes. It was argued that these discoveries corroborate the reliability thesis about formal and phenomenological memory markers.

References

Bernecker, Sven. *Memory: a philosophical study*. Oxford: Oxford University Press, 2010.

Carruthers, Peter. "How We Know Our Own Minds: *The Relationship between Mindreading and Metacognition.*" Behavioral and Brain Sciences 32, no. 2 (2009): 121–38. https://doi.org/10.1017/s0140525x09000545.

Dokic, Jérôme. "Feeling the Past: A Two-Tiered Account of Episodic Memory." *Review of Philosophy and Psychology 5, no. 3 (2014): 413–26. https://doi.org/10.1007/s13164-014-0183-6.*

Goldman, Alvin. *Simulating Minds: the philosophy, psychology and neuroscience of mindreading*. Oxford: Oxford University Press, 2006.

Kahneman, Daniel. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux, 2011.

Koriat, Asher. "The contentive confidence in one's knowledge and judgements: some metatheoretical considerations". *Foundations of metacognition*, edited by Michael Beran, Johannes Brandl et al., 213 - 233. Oxford: Oxford University Press, 2012.

Michaelian, Kourken. *Mental Time Travel: episodic memory and our knowledge of the personal past*. Cambridge, MA: The MIT Press, 2016.

Michaelian, Kourken; Perrin, Denis. "Memory as mental time travel". *The Routledge Handbook of the Philosophy of Memory*, edited by Sven Bernecker and Kourken Michaelian, 228 – 239, New York: Routledge, 2017.

Perner, Josef. "MiniMeta: in search of minimal criteria for metacognition". *Foundations of metacognition*, edited by Michael Beran, Johannes Brandl et al., 94 – 116, Oxford: Oxford University Press, 2012.

Proust, Joëlle. *The Philosophy of Metacognition: Mental Agency and Self-awareness* Oxford: Oxford University Press, 2013.

Schacter, Daniel; Addis, Donna. "The cognitive neuroscience of constructive memory: remembering the past and imagining the future". *Philosophical Transactions of the Royal Society of London.* Series B, Biological Sciences 362, 1481 (2007): 773 – 786. `https://doi.org/10.1098/rstb.2007.2087`.

Meaning in Gibberish: In Defence of Deep Bullshit

Tom Burdge*

University of St Andrews

Abstract In this essay I examine G. A. Cohen's notion of deep bullshit; I provide a counterexample to the often-implicit belief that deep bullshit is always bad, and unphilosophical. Section 1's outline of deep bullshit includes an important criterion for being deep bullshit which philosophers often leave implicit; deep bullshit is a bad and undesirable phenomenon that we should root out. Section 2 examines whether the *kōans* of Chan Buddhists were deep bullshit. In this section I argue they were; not only do they fit all modern definitions of deep bullshit, the Chan teachers were intentional deep bullshitters. In section 3 I argue we should not see the deep bullshit of Chan Buddhism as bad and un-philosophical; in Chan, through deep bullshit came philosophical inquiry. Section 4 responds to the Cohenian position, which holds any text which is "suggestive" is not deep bullshit; a Cohenian could claim Chan's *kōans* are suggestive, and so are not deep bullshit. I criticise this position by arguing that since philosophers of deep bullshit categorise Sokal's spoof article as deep bullshit, they must also categorise Chan *kōans* as deep bullshit. In section 5 I argue most allegations of deep bullshit are likely to be epistemic trespass. In section 6 I make recommendations for how to avoid trespassing deep bullshit allegations in future.

1   Deep Bullshit

The philosophical discussion of bullshit originates in Harry Frankfurt's work. Frankfurt identified Bullshit as a form of deceit, which was distinct from lying. Upon his account, when one lies one has the intention to convince the listener of a proposition which is not true, whereas when one bullshits one has no interest in the truth whatsoever.[1]

---

*Tom is a final year philosophy undergraduate and Laidlaw scholar at the University of St Andrews. Tom is particularly interested in philosophy of nonsense and Buddhist philosophy. He is the host of the Buddhist Philosophy Podcast, which you can find at `www.buddhistphilosophy.co.uk/listen`.

1. Harry G. Frankfurt, On Bullshit (Princeton, New Jersey: Princeton University Press, 2005), 52-4.

Cohen[2] mainly agreed with Frankfurt's definition of bullshit but thought he could identify bullshit of a different kind. Whereas Frankfurt's bullshit was identifiable from the intentions of the speaker, Cohen's bullshit is identifiable from the speakers' expressions themselves. Cohen called his new category "deep bullshit"; he believed it was especially prevalent in academia.

Cohen argued a condition for being deep bullshit is that a statement has "unclarifiable unclarity"; the statement is obscure and cannot be unobscured. Cohen argued that for unclear statements "adding or subtracting (if it has one) a negation sign from a text makes no difference to its level of plausibility".[3] If we could put a negation sign on Heidegger's definition of *dasein*, or Russell's definition of a set, and leave it equally plausible, then it would be unclear. If the statement can't be clarified to a sufficient level to change the symmetry of plausibility in negation, then it is deep bullshit. Below I call this the negation condition for deep bullshit; Cohen saw it as a sufficient condition.

Cappelen and Dever also gave an account of deep bullshit; they argued deep bullshit is "nonsense or gibberish"; the expression has no meaning, although it may appear to.[4] Non-exhaustive examples may include instances of one's words referring to things which don't exist, or one's expressions failing to satisfy the necessary conditions for meaningfulness (which they decline to give).[5]

A final, very important but often implicit, aspect of all bullshit is that it's bad. Frankfurt originally wanted to explore a phrase already used in everyday language, which we take as a bad quality of a statement, and investigate what this bad quality is. He claimed we use the word "bullshit" to describe the statements of bullshitting speakers because we find their statements, like excrement, "so repulsive".[6] Cohen wrote of how allegations of bullshit serve to "stigmatize" a text, and urged us to "conduct a struggle against" bullshit in the academy.[7] This is also true of Cappelen and Dever's recent chapter on bullshit, they call for a crusade to "root (deep bullshit) out".[8] The use of these words about bullshit are illuminating because of how they look down on statements which they classify as (deep) bullshit.

Unlike with Frankfurt's bullshit, it is possible to deep bullshit unintentionally. The unintentional deep bullshitter is probably gullible, but they are not morally blameworthy. It is the charlatan who, through "speak(ing) meaninglessly on purpose", intends to "deceive" by presenting their expressions as if they have meaning.[9] This inten-

---

2. Gerald Allan Cohen, "Deeper Into Bullshit", in The Contours of Agency: Essays on Themes from Harry Frankfurt, ed. Sarah Buss and Lee Overton (Cambridge, Massachussetts: The MIT Press, 2011).

3. Gerald Allan Cohen, "Deeper Into Bullshit", 333.

4. Herman Cappelen and Josh Dever, Bad Language (Oxford, UK: Oxford University Press, 2019), 60.

5. *Ibid*, 63-4, 67-9.

6. Frankfurt, On Bullshit, 43-4.

7. Cohen, "Deeper Into Bullshit", 335.

8. Cap.& Dev., Bad Language, 65.

9. *Ibid*, 64-5; Alan D. Sokal and Jean Bricmont, Fashionable Nonsense: Postmodern Intellectuals' Ab-

tional deception, on Cappelen and Dever's view, gives the charlatan a higher amount of reprehensibility than the unintentional deep bullshitter.

2   Deep Bullshit in Kōans

Chan Buddhism is a Chinese school of Buddhism. *Kōans* emerged in Chan from the 11th century. A *kōan* is a very concise story which is meant to help the reader achieve progress in their spiritual practice. *Kōans* commonly involved a question and answer between a student and a teacher, with the teacher's response expressing a great truth. In this section I argue many *kōans* are deep bullshit.

Cohen's negation condition for deep bullshit is especially true of a certain type of *kōan.* In certain *kōans* the student asks the teacher "what is Buddha?". In *kōans* 30, 33, and 34 of a single volume collection of *kōans* the teachers respond: "the very mind is Buddha", "no mind, no Buddha", and "mind is not Buddha".[10] The contradiction here is evident, but at no point in this collection of *kōans* is it suggested that the logically contradictory answers are incompatible; Chan holds the negation of one answer to "what is Buddha?" to be equally as plausible as the non-negated form. Therefore, it seems clear that Chan *kōans* fit Cohen's unclarifiable unclarity sufficient condition, and a Cohenian would categorise Chan *kōans* as deep bullshit.[11]

Would Cappelen and Dever categorise the "What is Buddha?" *kōans* as deep bullshit? It is difficult to speak on their behalf given their refusal to give necessary or sufficient conditions. But there are certainly examples which interpreters hold to be nonsense, or gibberish. For example, in another "what is Buddha?" *kōan* the teacher responds "*Masagin*! (Three pounds of flax!)".[12] Certainly there is a metaphorical element to the expression here, but Yamada's commentary on the case stresses that a central part of any interpretation should focus on the sound, the gibberish noise, with nothing else.[13] You could substitute *Masagin* for "whack (hitting the table)!" or just "*Ma!*", but the response should involve a nonsense element.

---

use of Science (New York: St. Martins Press, 1999), 5. Sokal and Bricmont coined "charlatan" for the intentional bullshitter; Sokal is introduced in §4.

10. Yamada Kōun, The Gateless Gate: the Classic Book of Zen Koans, 2nd ed. (Somerville, MA, USA: Wisdom Publications, 2015), 148, 161, 165.

11. There are attempts to clarify Chan Buddhist metaphysics. Although, even on these readings, many key Chan Buddhist concepts behind kōans such as "transcending duality" remain "not, of course, entirely clear''. Graham Priest, "Enlightenment'', in The Fifth Corner of Four: an Essay on Buddhist Metaphysics and the Catuskoti, (Oxford: Oxford University Press, 2018), 141. See chapter 9 of this work for as clear a reading of Chan metaphysics of enlightenment as is likely possible. For any reader still convinced that all kōans can be clarified, this essay can be read as saying even if it were the case that kōans were unclarifiable, it would not follow that they were something to be rooted out. I am grateful to an anonymous peer reviewer for this clarification.

12. Yamada, The Gateless Gate, 89.

13. *Ibid*, 90.

Reflective readers will already be unhappy with Cohen, and Cappelen and Dever's accounts of deep bullshit. Cohen's definition itself is unclear; what does it mean to be "plausible"? This is quite a dubious notion for a number of reasons.[14] Cappelen and Dever's account is also quite underdeveloped. For instance, when talking about French psycho-analyst and philosopher Jacques Lacan, Cappelen and Dever are very confident in claiming "much of what Lacan said was meaningless", but they decline to give necessary and sufficient conditions for a statement to have meaning.[15]

I'm sure there are responses to the above criticisms, but to an extent this doesn't matter. "Poking fun" and deprecation of peoples' life works is very common in the Chan tradition, sometimes it almost seems as if the higher the mockery, the higher the praise.[16] For example, Bodhidharma (the monk who brought Buddhism to China) is called a "broken toothed barbarian", but Mumon, the compiler of this work, insists that what initially sounds insulting (not because of any mistranslation) has "friendly overtones".[17] In this sense, when one says that Chan philosophy is "bullshit" it sounds like one is affirming, even grasping the true message, of the tradition.

I believe the Chan Buddhists were intentional deep bullshitters. To me, despite the weaknesses of Cappelen and Dever's account, deep bullshit as some sort of profound-sounding gibberish is still intuitive. For example, I would be more than happy to refer to something that comes from a New Age Bullshit generator as deep bullshit.[18] In Froese's interpretation of Chan, she argues for an interpretation to the effect that they intentionally created their own deep bullshit.[19] There is a story of a Chan teacher who speaks to a congregation of the "great mystery", claiming most do not appreciate its "application". Yang Shan, a student, asks for clarification on these words, and is promptly kicked in the face by the monastery's abbot, at which the teacher laughs.[20] Froese takes this story as the teacher baiting the students into seeking intellectual pursuit through "grandiose words"; the words used are intentionally profound-sounding, but are in fact deep bullshit.

3   The Meaning of the Gibberish

Cohen wanted to formulate deep bullshit independently from intentions, but, as in the case of charlatans, if the author aims at deep bullshit it is likely that they succeed. In

---

14. I am indebted to comments from Walter Pedriali in an email exchange for this point.

15. Cap.& Dev., Bad Language, 70-1.

16. Katrin Froese, Why Can't Philosophers Laugh? (Cham, Switzerland: Palgrave Macmillan, 2017), 174.

17. Yamada, The Gateless Gate, 194-7.

18. See: https://sebpearce.com/bullshit/

19. Froese Why Can't Philosophers Laugh?, 178.

20. Tao-Yuan and Sohaku Ogata, The Transmission of the Lamp: Early Masters (Wolfeboro, NH: Longwood Academic, 1990), 300.

my view, the Chan Buddhists certainly intended for their words to be deep bullshit. But I wouldn't call the Chan teachers charlatans, because I don't see their discourse as bad and un-philosophical. I justify this through answering why Yang Shan is kicked.

On the one hand, Yang Shan is refused clarification; the language the Chan teachers are using embraces contradiction and unclarity as a way to expose what they believe is the inevitable paradox and bullshit in all language; their statements draw attention to a universal unclarity they think is present in language.[21] On the other hand, the kick is Yang Shan's clarification; you will find what the teacher means outside of language, through experience of day to day sensations like the pain of a kick in the face.

To me, these interpretations of the Chan teachers gives them a philosophical position. Chan has a position on the nature of all language; it is all gibberish. Phrases like "the sound of one hand clapping" serve to illustrate this but are no more gibberish than the rest of language. The Pyrhonnian sceptics held a similarly unoptimistic position about the nature of all justification; if the Pyrhonnians' position doesn't exclude them from being philosophers, it seems difficult to non-chauvinistically exclude Chan Buddhism for their comparable positions.

Furthermore, Chan takes a normative position on the best course of action to accumulate virtue; it is awareness of everyday experience. They thought that the contradiction in all statements entailed that we should focus on what is immediately aware to us in perception. Their normative ethics is like any other in philosophy. They use reasoning to find the best course of action; the only difference is that with their use of logic they also reject logic as we know it.

Each *kōan*'s idiosyncratic response, like a thought experiment, functions as a philosophical tool; it serves to illustrate Chan's broad theses about truth. The "what is Buddha?" *kōan*s' meaning is that there is no meaning, or at least no non-contradictory meaning; in this sense, there is meaning to the gibberish. There is just "whack!" The kōan functions as a pedagogical tool to frustrate, and then gradually remove attachment to language. Through tying our brain into knots with its deliberate nonsense, the kōan's message is to highlight the unsatisfactoriness of language as a tool for seeking happiness.[22] This is certainly obscure, unclear, and even deep bullshit, but should we root it out? Is it not worthy of the name philosophy?

We might complain that any account of meaning which includes gibberish is unsatisfactory, but do we already have sufficient handle on the notion of meaningfulness to exclude gibberish? Without assuming that the methodology that has been useful to analytic philosophers for **their own** purposes and discussions for the last 100 or so years is superior and philosophy-defining, can we justifiably say that the Chan approach is bad and unphilosophical?

Some opponents to my view might admit that the Chan teachers were not being

---

21. Froese, Why Can't Philosophers Laugh?, 22, 182-3.
22. *Ibid*, 18.; Useful quote from this page: "Sense becomes nonsense; nonsense becomes sense".

unphilosophical with their *kōan*s and more general methods of inquiry, but the opponents would maintain the claim that what the Chan teachers were doing was **bad**. Indeed, claiming the Chan inquiry was good philosophy is a far stronger claim than Chan philosophy was philosophical in any sense, and the two are separable. One might admit that the Chan Buddhists aimed to answer similar questions to good philosophers but used a poor methodology to do so. This section has made a very tentative attempt at defending both claims. I hope that this section provides little doubt that the Chan Buddhists were doing philosophy in some form; they made similar claims to others whom we call philosophers. A much longer essay would be needed to indisputably defend Chan inquiry as **good** philosophy. Certainly, if I am correct that there is good Cohenian bullshit, then further work needs to be done to distinguish the good from the bad.

4   Productive Suggestiveness is not an Anti-necessary Condition

One might argue that, despite the arguments in Section 2, Chan *kōan*s are not deep bullshit. Cohen appealed to "suggestiveness" to classify "good poetry" that is still unclear/gibberish, as not deep bullshit. Cohen wanted to formulate his deep bullshit independently from the speaker's intentions, preventing him from arguing that poetry is not deep bullshit because it is designated *as* poetry by the speaker. Instead, he argued an unclarifiable text could be "valuable", i.e. not deep bullshit which is bad, "because of its suggestiveness". Through the multiple interpretations that we can draw from the text it can "stimulate thought" to give us valuable insights.[23]

Cohen himself, in a reference, gave a very brief suggestion about what his own views might have been on Chan Buddhism. He allows "the unclarifiable may be productively suggestive" but does not agree with Yu-Lan's claim that Chinese philosophy's common lack of "articulateness" (presumably, close enough to be a synonym for clarity for Cohen's use) is compensated for by Chinese philosophy's "almost boundless" suggestiveness.[24]

But let's move past Cohen's personal biases and assume that Chan Buddhism is "productively suggestive" (as said in section 3, to show that Chan Buddhism is good philosophy would need a much longer essay). In the context of philosophy, productive suggestiveness is the ability of a text to "stimulate thought" in a manner which is productive in helping us reach valuable philosophical insights. Cohen held productive suggestiveness as an anti-necessary condition (the negation of the condition

---

23. Cohen, "Deeper Into Bullshit", 333-4.
24. *Ibid*, 334.; Fung Yu-Lan, A Short History of Chinese Philosophy, ed. Derek Bodde (London: Macmillan, 1960), 12.

is necessary) for being deep bullshit that overrode his unclarifiable unclarity sufficient condition; if an unclarifiable text is productively suggestive, then it is not deep bullshit.

I would agree with Cohen that the productive suggestiveness of Chan *kōan*s, assuming that they are productively suggestive, is what makes them philosophically valuable. However, I would maintain that Chan *kōan*s are often deep bullshit; I just wouldn't condemn them, and label them as non-philosophy because of this. We can tell Chan *kōan*s are deep bullshit because Froese's interpretation of the Yang Shan kicking case tells us the Chan Buddhists **aimed** at deep bullshit.[25] Sokal made a spoof article which Cohen accepted as "deliberate bullshit"; the piece was "self-condemning" in that it was presented as non-deep bullshit at first, and when it fooled people as sincere it was supposed to illustrate a point about the nature of philosophical inquiry (namely, that lots of philosophy conducted on the continent is deep bullshit).[26] Just as with Sokal, Chan Buddhists aimed to tell us something about the nature of philosophical inquiry through producing a piece of self-condemning philosophy; the meaning was that the statements they were making inevitably included some deep bullshit, which makes a broader point about philosophical inquiry. What this shows is that, even with their productive suggestiveness, Chan Buddhist *kōan*s are still deep bullshit, because they intentionally embrace deep bullshit just like Sokal's article. It is also true that, like Sokal's article, it is perfectly possible for the *kōan*s to be insightful despite being deep bullshit.

A Cohenian might complain that the response in the previous paragraph ignores Cohen's formulation of deep bullshit as independent from "intentional encasement" i.e. the intentions of the speaker.[27] Cohen admitted Sokal's article was deep bullshit, but not because of Sokal's intention to create deep bullshit. There were some other satisfied conditions which made Sokal's article deep bullshit, including the article's unclarifiable unclarity. That Sokal was open about his intent to create deep bullshit is helpful in identifying that the article is deep bullshit, but the intent is not what makes the article deep bullshit.

To me this doesn't seem right; the reason Sokal and Cohen saw Sokal's article as "really" deep bullshit is because of Sokal's devious intention to create deep bullshit. But, even if it is true Sokal's article was deep bullshit because of an intention-independent sufficient condition for deep bullshit being satisfied, I would challenge objectors to identify a condition which can be found in Sokal's article that classifies it as deep bullshit which is not present in the teacher's statement in the Yang Shan case. We might claim Sokal's article was deep bullshit because it was obscure, unclear, and

---

25. This doesn't make the kōans deep bullshit on the accounts of deep bullshit we are working with. However, this does provide a good additional indicator; if the speaker aimed at deep bullshit, and upon analysis their statements seem to be a profound-sounding nonsense, then their statements are most likely deep bullshit. I am grateful to an anonymous referee for this point.

26. Alan D. Sokal, "Transgressing the Boundaries: Toward a Transformative Hermeneutics of Quantum Gravity'', Social Text, no. 46/47 (1996): , 217-252, https://doi.org/10.2307/466856.

27. Cohen, "Deeper Into Bullshit", 7.

unverifiable; these are all true in the Yang Shan case.

Cohen's classification of Sokal's article as the archetype of that which aims to be deep bullshit and succeeds, and the similarity of productively suggestive argumentation from Sokal's intentional deep bullshit to Chan's intentional deep bullshit is noticeable.[28] Therefore, a Cohenian claim that the productive suggestiveness of Chan excludes it from being deep bullshit cannot work; assuming Chan is productively suggestive, it is so in the same way as Sokal's article (which Cohen is happy to call deep bullshit).

In the current and previous sections, my goal has been to question the narrative that deep bullshit is not always negative with examples from Chan Buddhist philosophy. Another aim was to suggest that nonsense might have some meaning, but this is a far more ambitious aim that I do not claim to have proven.

In the following sections, I want to reflect on why deep bullshit allegations have arisen.[29] In section 5, I analyse why deep bullshit allegations emerge; I argue that epistemic trespass is often the most likely explanation for deep bullshit allegations. In section 6, I make some recommendations for avoiding false, trespassing, deep bullshit allegations.

5 The Best Explanation for Allegations of Deep Bullshit is Epistemic Trespass

Epistemic trespass occurs when an expert has competence to make judgements in one field, but they speak about another field where they lack competence. With so-called public intellectuals, it is easy to think of cases of this. Richard Dawkins is one example. Dawkins is certainly an expert on evolutionary biology, but he has also published books and given talks about religion. Philosophers of religion, of all beliefs about the existence of God(s), have accused Dawkins of being very uncharitable to his opponents and ignoring the genuine issues.[30] Despite his lack of expertise in this domain, Dawkins speaks with as much confidence as he would if he were talking about evolutionary biology.

Why do deep bullshit allegations emerge? For most instances, I think there are two explanations that are most plausible:

---

28. Cap.& Dev. Bad Language, 64. Cappelen and Dever call Sokal's work "intentional gibberish" so they may not be so susceptible here Although, my reading of Cappelen and Dever would interpret them as being happy with calling Sokal's spoof deep bullshit.

29. Katherine Hawley, "Identity and Indiscernibility'', Mind 118, no. 469 (January 2009): , 101-119, https://doi.org/10.1093/mind/fzn153. This is strongly influenced by Katherine Hawley's change of direction in the final section of her paper on the Identity of Indiscernibles.

30. Nathan Ballantyne, "Epistemic Trespassing'', Mind 128, no. 510 (December 2018): , 367, https://doi.org/10.1093/mind/fzx042.

1. The allegation is correct. The accused is producing deep bullshit; possibly intentionally. [31]

2. The allegation is incorrect. The accused is not producing deep bullshit. The accuser has committed epistemic trespass.

There are other plausible explanations. Perhaps the accuser is incorrect but there is no trespass, maybe they just want to discredit their dialectical opponent and are committing (Frankfurt's) bullshit, but I imagine these are relatively rare cases.

I don't want to make a systematic survey of all allegations of deep bullshit, and claim most of them are instances of epistemic trespass. Instead, I want to suggest that in most situations the allegation is best explained by epistemic trespass, rather than genuine deep bullshit. But, to illustrate the point, Sokal and Bricmont's high profile allegations against Lacan are useful.[32] Lacan was a prolific French psycho-analyst, who has been highly influential for continental philosophy.[33] Sokal and Bricmont argued Lacan created deep bullshit; they took particular issue with his claims which make use of mathematical language. For instance, they expressed disdain for Lacan's equating "the erectile organ. . . to the $\sqrt{-1}$ ".[34]

Perhaps 1. is true. Lacan was competent enough to achieve highly in academia, so he and his most intelligent successors probably also figured out that this is nonsense. They have gone along with the bullshit; they are knowing charlatans. The charlatans have conned philosophers, anthropologists, sociologists, Marxists, and those in many other fields. They have also fooled the students who study under them and who pay lots of money to be taught by them (although, not if they are being taught at a public French university).

Option 1. would be the most likely explanation if:

(i) The accuser(s) are experts in the field that they are accusing someone in.

(ii) The accused are talking about a domain outside of their own expertise, making the accused more likely to be trespassing.

---

31. Despite my non-pejorative use of "deep bullshit" to refer to discussions in Chan philosophy above, I have not seen any other deep bullshit attributions that are non-accusatory, and non-pejorative.

32. Sokal and Bricmont don't use the phrase "deep bullshit", but Cappelen and Dever, and Cohen, took them as showing that many "French thinkers" wrote deep bullshit.

33. Adrian Johnston "Jacques Lacan" Stanford Encyclopedia of Philosophy, ed. Edward N. Zalta (2018).

34. Sokal & Bricmont, Fashionable Nonsense, 27.; Jacques Lacan, in Ecrits: a Selection., trans. Alan Sheridan (New York: W.W. Norton, 1997), 292-325.; For a defence of Lacan's statement see Arkady Plotnitsky, The Knowable and the Unknowable: Modern Science, Nonclassical Tought, and the "Two Cultures" (Ann Arbor: The University of Michigan Press, 2005). Arkady Plotnitsky, "On Lacan and Mathematics", Oeuvres & Critiques XXXIV 2 (2009): , 143-162.

(iii) There is a consensus, within the field that the accused's work is part of, that the accused's work is deep bullshit.

None of these are the case for Sokal and Bricmont's allegations against Lacan's statements that Sokal and Bricmont quote.[35]

i. does not apply. Sokal and Bricmont seemed to be aware they were writing about thinkers whom they were not experts on.[36] The implication seems to be that when in their allegations Lacan wrote 'nonsense' they limited themselves to areas where Lacan made use of mathematical concepts, on which they are experts. But in Sokal and Bricmont's treatment of Lacan this is blatantly not the case. For instance, when Lacan re-appropriates the mathematical concept 'compactness' for his psycho-analysis and gives a definition of compactness in the psycho-analytic sense, they argue Lacan's definition of "compactness is not just false: it is gibberish".[37] They make a similar claim that the concept "space of *jouissance* in psycho-analysis" is "ill defined".[38] In both cases, what most indicates their lack of expertise is their failure to provide **any** justification whatsoever for why, when treating the term in a psycho-analytic context and not a mathematical one, the concepts they criticise are "nonsense". They seem to believe the texts they quote "speak for themselves" as nonsense, but this is merely a way of appealing to the intuitions of readers already sympathetic to Sokal and Bricmont's position rather than an example of any competence with the material they criticise.[39] When writing about both 'compactness' and the 'space of *jouissance*' Sokal and Bricmont stray into talking about the psychoanalytic, non-scientific aspects of Lacan's work which they readily admit they are "not competent to judge".[40]

ii. does not apply either. It might initially seem like he was trespassing; equivocating the "erectile organ" (whatever that may be) to $\sqrt{-1}$ does initially seem like a fantastical claim about mathematics. Mathematics, in the sense that we usually think of it, is a domain that Lacan as a psycho-analyst was not an expert in. But Lacan was not actually writing about the same $\sqrt{-1}$ as Sokal and Bricmont. Under Plotnitsky's interpretation of Lacan there are two types of $\sqrt{-1}$.[41] The first is (M) $\sqrt{-1}$, the mathematics we immediately think of, on which Sokal and Bricmont have a good level of expertise, and on which Lacan was quiet. The second is (L) $\sqrt{-1}$, Lacan's very broadly analogous mathematics, which he invented. Lacan's "mathematics" was only strictly mathematical in the sense that it made an analogy between psycho-analytic phenomena and complex numbers.

---

35. Sokal & Bricmont, Fashionable Nonsense, 18-37.
36. *Ibid*, 7.
37. *Ibid*, 22-3.
38. *Ibid*, 9.
39. *Ibid*, 37.
40. *Ibid*, 7. 'It goes without saying that we are not competent to judge the non-scientific aspects of these authors' work'.
41. Plotnitsky, The Knowable and the Unknowable, 113.

Lacan argued truths captured by psycho-analytic concepts are extremely difficult, perhaps impossible, to visualise and conceptualise.[42] Lacan compared his psycho-analytic concepts to complex numbers; a similarly elusive phenomena which, when we signify them, we signify not the thing itself, but an image of the thing.[43] Lacan constructed a much more complex (L) mathematical system than this paper can outline, but his claims always remained psycho-analytic while employing only very broadly mathematical analogies.[44] In this light, Lacan would be the foremost expert to talk about (L) mathematics; he invented the system.

Finally, iii. does not hold. there is not a consensus among those who are experts on Lacan that what Lacan said was deep bullshit. Therefore, 1. seems an unlikely option in this case.

The alternative option is 2; the accuser has committed epistemic trespass. The expert in their field, as the high-profile cases usually are or else they wouldn't have enough social capital for anyone to listen to them, has strayed from their own domain to talk about something which they wrongly believe they have the proficiency to speak on.

I think that, in most cases where allegations of deep bullshit have emerged, 2. is the most reasonable explanation. In the allegations of deep bullshit I have mentioned above, none of i-iii apply. I think attention to the absence of i. in the accusers gives the most insightful explanation of why the allegations are there. Investigating further, the accuser initially might **seem** to be accusing within their own domain. For example:

- Sokal and Bricmont **initially seem**[45] to be criticising Lacan's interpretation of something they are experts on; mathematics.

- Cappelen and Dever are experts on philosophy, and while they call Lacan a "thinker" rather than a philosopher, it would be easy to think that their philosophical training makes them an authority on Lacan given his influence on (continental) philosophical fields.[46]

- Cohen was no doubt an expert on Marxism, but his readings on French Althusserianism as a youth do not make him an expert on it; he helped found Analytical Marxism, an entirely different research project.

---

42. It is beyond the scope of this paper to defend this claim.

43. Plotnitsky "On Lacan and Mathematics''. 151-2.

44. Sokal and Bricmont's complaints 'misuse' of mathematics in analogies are separable from their claims about nonsense, with which this essay is concerned. They make the distinction: Sokal & Bricmont, Fashionable Nonsense, 4-5.

45. I am indebted to a series of lectures on paradoxes during Michaelmas 2019 by Patrick Greenough at the University of St Andrews for the phrasing "initially seems".

46. It is also possible that they are appealing to the authority of Sokal and Bricmont, who confirm their bias.

For all the above, it initially seems as if the accusers are experts in the domain within which they accuse, but this is in fact not true. This lends weight to the theory that deep bullshit accusations are often instances of epistemic trespass due to overzealous transfer. Overzealous transfer is when agents transfer their skills to another context, but their skills are inappropriate.[47] Due to the initial seeming proximity of the fields the accusers are experts in to the one in which they accuse, they falsely think that their expert-domain skills will transfer well to the accusing-domain. In Sokal and Bricmont's case, this mistake is most obvious.

I hope that this section has successfully suggested that deep bullshit allegations are often instances of epistemic trespass due to overzealous transfer, rather than the truth of a conspiracy theory in French academia. Ballantyne claims that examples of epistemic trespass allegations are widespread, and I hope that my recommendations in the following section can help prevent further trespassing deep bullshit accusations.[48]

6   Some Recommendations

To avoid instances of future trespassing deep bullshit allegations I make the following three recommendations:

First, have greater epistemic modesty. Before making any allegation, ask yourself: do I really have enough expertise to reliably make this deep bullshit allegation? Do I have a good track record in this domain, or does it just initially seem that I do? If the domain is outside of your own expertise, ask: do I have enough cross-field expertise to make this allegation?[49]

Second, don't use generics around instances of deep bullshit. In the same volume as their chapter on bullshit, Cappelen and Dever have a chapter dedicated to pointing out the dangers of generics in leading us to cognitive error.[50] A generic is, roughly, a general claim about a certain kind that is vague as to how strong it is between an existential and universal claim e.g. 'cats aren't loving' or 'French philosophers come up with lots of deep bullshit'.[51] Despite the warnings, Cappelen and Dever are perfectly happy to use relatively vague quantifying determiners in their claims, such as "much of what Lacan said was bullshit".[52] This makes their claims very difficult to falsify; if I were to give them some quotes from Lacan, they could simply say 'oh, no I wasn't talking about those cases'.

---

47. Ballantyne, "Epistemic Trespassing'', 385.
48. *Ibid*, 369.
49. *Ibid*.
50. Cap.& Dev., Bad Language, 126-43.
51. The latter is a quote from a lecture given by Professor Cappelen during Candlemas semester 2018 at the University of St Andrews. It is more specific than the cats example, but still vague as to the strength of the claim.
52. Cap.& Dev. Bad Language, 66, 71.

To me, the third recommendation is most important of all, as the first two should have always been fairly obvious. I believe the splits that exist within philosophy have made the deep bullshit allegations more acceptable. In addition to the overzealous transfer explanation, I think allegations of deep bullshit represent a chauvinistic attitude toward other philosophical traditions. Whereas Chan mockery toward Bodhidharma, because he was an Indian "barbarian" who didn't know Chinese, was playful mockery, philosophical close-mindedness is often entirely serious.[53] An example from a lecture by Professor Cappelen should be revealing here. Cappelen divulged an anecdote about a friend who studied Lacan prior to starting "real philosophy", before Cappelen quickly corrected himself as having "mis-spoken".[54] The French thinker conspiracy, like many other conspiracy theories, has emerged in large part from suspicion toward those belonging to a different social group.

A disparaging attitude towards other philosophical traditions is not unique to the analytic school. Jay Garfield, in an account of how he became interested in Buddhist Philosophy, writes about how the late director of the Institute of Buddhist Dialectics, Gen Lobsang Gyatso, told him western philosophy was "shallow and materialistic" when they first met.[55] Fortunately, Garfield's teachings on Kant's *Critique of Pure Reason* allowed Gyatso to notice the trespass.[56] Similarly, Jacques Bouveresse, an analytic philosopher at a French university, wrote in 1983 about how fellow French philosophers dismissed analytic philosophy as useless and "in the process of dying" because it was only concerned with logic. These cases were clearly trespass, given the same individuals often thought Wittgenstein was a logical positivist.[57]

The split in philosophy between analytic and continental, and between Eurocentric and non-European, non-English speaking, and/or non-Eurocentric philosophies exists. Strassfield's work shows that the split between continental and analytic in America was because of a "brahmin caste" of elites who ensured analytic philosophy became dominant.[58] Garfield and Van Norden claim that the split between western and non-western philosophy has emerged as a legacy of colonial euro-centrism.[59] Hence, historically speaking, neither split was as a result of the superiority of a specific methodology.

---

53. See §4 for this example.

54. During a lecture given by Cappelen during Candlemas semester 2018 at the University of St Andrews.

55. Jay L Garfield *Practicing without a License and Making Trouble along the way: My Life in Buddhist Studies*. (2018), 5. https://jaygarfield.files.wordpress.com/2018/10/practicing-without-a-license.pdf

56. I would like to stress that Gyatso's action is in no way comparable to the injustices which emerge from Eurocentric philosophy departments, merely that the disparagement is commonplace.

57. Jacques Bouveresse, "Why I Am so Very UnFrench", in Philosophy in France Today, ed. Alan Montefiore (Cambridge, UK: Cambridge U., 1983), 10-11, 13.

58. Jonathan Strassfeld, "American Divide: The Making Of 'Continental' Philosophy", Modern Intellectual History, 2018, 1, https://doi.org/10.1017/s1479244318000513.

59. Jay L Garfield and Bryan W Van Norden, "If Philosophy Won't Diversify, Let's Call it What it Really Is" The Stone (The New York Times, May 11, 2016), https://www.nytimes.com/2016/05/11/opinion/if-philosophy-wont-diversify-lets-call-it-what-it-really-is.html.

Ballantyne calls for fields with potential for trespass to "rub shoulders", but I hardly feel that this is a sufficiently high-reaching recommendation to correct the splits in philosophy which have led to trespass susceptibility.[60]

In 2020 at the University of Edinburgh I presented an early draft of this paper.[61] Before the talk, someone approached me and asked why I was presenting on Buddhism and philosophy, which they told me were very different things. Afterward, the same person told me kōans are poetry and not philosophy, of course without giving me any justification whatsoever when I asked them why they thought that. This was an instance of boundary policing; a claim that a paper is "not philosophy" rather than a genuine engagement which critiques the paper's arguments.[62] To me, it seems unlikely that genuine engagement with the positions presented will miraculously occur if we merely encourage philosophical shoulder rubbing between philosophers of differing traditions. To correct this, my final recommendation is more of a demand: we need an ambitious policy of diversification of philosophy departments' teaching and research interests, and a more general inclusivity of marginalised voices within philosophy.

7   Conclusion

In this paper I have argued deep bullshit is not always a dirty, unworthy phenomenon which we should root out of philosophy. In section 1 I quickly outlined Frankfurt's account of bullshit, before moving onto Cohen's deep bullshit. I highlighted the fact that all philosophers who have written about deep bullshit argue it is always a bad occurrence. In section 2 I introduced kōans and I argued some kōans are deep bullshit. However, in section 3 I argued these kōans are not a negative phenomenon which we should root out of philosophy. In doing so I argued, contra the assumptions of Cohen and Cappelen and Dever, deep bullshit is not always a bad occurrence. In section 4 I responded to a potential objection from a Cohenian which claims kōans are not deep bullshit, because they are productively suggestive. In response, I argue that because Cohenians take Sokal's paper as deep bullshit, they must categorise kōans as deep bullshit too. In section 5 I moved in a different direction to argue the best explanation for most occurrences of deep bullshit allegations are epistemic trespass, rather than genuine occurrences of deep bullshit or charlatanry. In section 6 I made some recommendations to avoid future spurious deep bullshit allegations. The most important recommendation argues that we need an ambitious project to transform philosophy

---

60. Ballantyne "Epistemic Trespassing", 388.

61. I had a great evening and my thanks go to the University of Edinburgh Philosophy and Buddhist societies for hosting me. I give this example because it is relevant, not because it represents my experience of the whole of the evening.

62. For a seminal paper on this phenomenon: Kristie Dotson, "How Is This Paper Philosophy?", Comparative Philosophy: An International Journal of Constructive Engagement of Distinct Approaches toward World Philosophy 3, no. 1 (2012), https://doi.org/10.31979/2151-6014(2012).030105.

departments so that trespassing deep bullshit allegations can no longer occur.[63]

References

Ballantyne, Nathan. "Epistemic Trespassing''. *Mind* 128, no. 510 (2018): 367–95. https://doi.org/10.1093/mind/fzx042.

Bouveresse, Jacques. "Why I Am so Very UnFrench''. Essay. In *Philosophy in France Today*, edited by Alan Montefiore, 10–30. Cambridge: Cambridge U.P., 1983.

Cohen, Gerald Allan. "Deeper Into Bullshit''. Essay. In *The Contours of Agency: Essays on Themes from Harry Frankfurt*, edited by Sarah Buss and Lee Overton, 321–44. Cambridge: The MIT Press, 2011.

Cappelen, Herman, and Josh Dever. *Bad Language*. Oxford: Oxford University Press, 2019.

Dotson, Kristie. "How Is This Paper Philosophy?" *Comparative Philosophy: An International Journal of Constructive Engagement of Distinct Approaches toward World Philosophy* 3, no. 1 (2012). https://doi.org/10.31979/2151-6014(2012).030105.

Frankfurt, Harry G. *On Bullshit*. Princeton, New Jersey: Princeton University Press, 2005.

Froese, Katrin. *Why Can't Philosophers Laugh?* Cham, Switzerland: Palgrave Macmillan, 2017.

Garfield, Jay L. (2018) *Practicing without a License and Making Trouble along the way: My Life in Buddhist Studies* available at: https://jaygarfield.files.wordpress.com/2018/10/practicing-without-a-license.pdf

Garfield, Jay L, and Bryan W Van Norden. "If Philosophy Won't Diversify, Let's Call It What It Really Is''. *The Stone*. The New York Times, May 11, 2016. https://www.nytimes.com/2016/05/11/opinion/if-philosophy-wont-diversify-lets-call-it-what-it-really-is.html.

Hawley, Katherine. "Identity and Indiscernibility''. *Mind* 118, no. 469 (2009): 101–19. https://doi.org/10.1093/mind/fzn153.

Johnston, Adrian. "Jacques Lacan''. *The Stanford Encyclopedia of Philosophy*. 2018. https://plato.stanford.edu/archives/fall2018/entries/lacan/

Yamada Kōun. *The Gateless Gate: the Classic Book of Zen Koans*. 2nd ed. Somerville, MA, USA: Wisdom Publications, 2015.

Plotnitsky, Arkady. *The Knowable and the Unknowable: Modern Science, Nonclassical Tought, and the "Two Cultures"*. Ann Arbor: The University of Michigan Press, 2005.

Plotnitsky, Arkady. "On Lacan and Mathematics''. *Oeuvres & Critiques XXXIV* 2 (2009): 143–62.

Priest, Graham. "Enlightenment''. Essay. In *The Fifth Corner of Four: an Essay on Buddhist Metaphysics and the Catuṣkoṭi*, 125–44. Oxford: Oxford University Press, 2018.

Lacan, Jacques. Essay. In *Ecrits: a Selection.*, translated by Alan Sheridan, 292–325. New York: W.W. Norton, 1997.

Sokal, Alan D. "Transgressing the Boundaries: Toward a Transformative Hermeneutics of Quantum Gravity''. *Social Text*, no. 46/47 (1996): 217. https://doi.org/10.2307/466856.

Sokal, Alan D., and Jean Bricmont. *Fashionable Nonsense: Postmodern Intellectuals' Abuse of Science*. New York: St. Martins Press, 1999.

Strassfeld, Jonathan. "American Divide: The Making Of 'Continental' Philosophy''. *Modern Intellectual History*, 2018, 1–34. https://doi.org/10.1017/s1479244318000513.

Tao-Yuan, and Sohaku Ogata. *The Transmission of the Lamp: Early Masters*. Wolfeboro, NH: Longwood Academic, 1990.

Yu-Lan, Fung. *A Short History of Chinese Philosophy*. Edited by Derek Bodde. London: Macmillan, 1960.

Can God Only Exist in the Present Moment?

FIONA COLLINS-TAYLOR*
University of York

Abstract   In this paper, I will argue that presentism is inconsistent with the belief that God exists in time. The defence of my argument will be split into two sections. The first will show that if God exists in time and presentism is true, there are some true propositions about the future. The second will use truthmaker theory to show that these propositions require an existing future entity to be made true. As the first section states that some propositions about the future are true, it is the case that some future entities exist. Hence, if God exists in time then some future entities exist. This conclusion goes directly against presentism's fundamental claim, proving it is inconsistent with the belief that God exists in time.

1   Introduction

Presentism is the view that only the present moment and objects within the present moment exist.[1] It is popular because it is a "common sense" approach to visualising time and is consistent with how we experience time, as changing from one moment to the next.[2] The importance of presentism has led to a wide discussion of its conjunction with God's relationship with time. Notable philosophers such as William Lane Craig have stated that God's existence in time (otherwise known as God's temporality) is consistent with presentism.[3] Alan R. Rhoda goes further in arguing that God's existence in time can solve the prominent truthmaker problem for presentism.[4] This paper

---

1. Ned Markosian, "A Defence of Presentism," in *Oxford Studies in Metaphysics*; Volume 1, ed. Dean W. Zimmerman (Oxford : Oxford University Press, 2004), 47.

2. Markosian, "A Defence of Presentism," 48.

3. William Lane Craig, *Time and Eternity: Exploring God's Relationship to Time* (Wheaton (ll.): Crossway Books, 2001).

4. Alan R. Rhoda, "Presentism, Truthmakers, And God," *Pacific Philosophical Quarterly* 90, no. 1 (2009): 41-62, https://doi.org/10.1111/j.1468-0114.2009.01328.x.

will show that the belief God exists in time is inconsistent with presentism. Therefore, philosophers should consider the possibility that God exists outside of time, reject truthmaker theory or abandon presentism.

I will present a unique defence which combines God's effect on the future with truthmaker theory. Following this introduction, I will summarise my argument, and its subsequent defence will be divided into sections 3 and 4. In section 3, I will prove the premise that if God exists in time then there are some true propositions about the future. Section 4 will be divided into two sub-sections. In section 4.1, I will use truthmaker theory to show that the propositions discussed in section 3 require an existing future entity to be made true. In section 4.2, I anticipate various objections which attempt to locate truthmakers in the present moment. These objections fail to supply truthmakers for the propositions being discussed because they can only be made true by the existence of the entities that they refer to. As section 3 shows that these propositions are true, it is the case that some future entities exist. Presentists state that future entities do not exist, hence the conclusion of this paper, that the belief that God exists in time is inconsistent with presentism. In other words, God cannot only exist in the present moment.

The rest of this introduction will be used to adapt God's attributes. The God of classical theism is described as an immutable and simple being who exists outside of time for all of eternity. If God exists in time and presentism is true, like every other existing entity which can only exist now, God only exists in the present moment.[5] Consequently, God's traditional attributes, specifically His immutability, omniscience and omnipotence, must be discarded or adapted to function within the confines of presentism.

Firstly, I discard the attribute of immutability. If God is immutable, He is unchanging. If God exists in time and presentism is true, He must change. An outcome of presentism is that propositions truth values change as time changes. For example, in the present moment the proposition 'that I am sitting down' is true. In an hour when I am standing this proposition will be false. God knows it is true in the present moment and will know it is false in the future. God's knowledge changes when propositions change their truth value. Therefore, God changes. I have discussed this attribute because a theist may be reluctant to discard it and argue that its loss is problematic. As it is an unavoidable outcome of God existing in time and presentism being true, I will not raise any objections.

The attribute of omniscience implies that God's knowledge is limitless. In this interpretation, God has experienced the past and is experiencing the present moment. God has knowledge of everything that has occurred in the past and has knowledge of everything that is occurring in the present moment.[6] God's knowledge of the future

---

5. Rhoda, "Presentism, Truthmakers, And God," 53.

6. *Ibid.*, 54.

is not as easily defined. It depends upon whether it is believed that the future is undetermined, determined or partly determined. Rhoda's view on the future is unclear. However, he does state that God can anticipate the future.[7] As I expect Rhoda's paper will be used as the main argument against my conclusion, I will maintain that God can anticipate the future.

Lastly, God's omnipotence can be redefined as the attribute of being all powerful within the present moment. God cannot directly act in the past or future because they do not exist. God can only act in the present moment and within the present moment His power is limitless.

2 Summary of the Argument

I will argue that presentism is inconsistent with the belief that God exists in time. A summary of my argument is presented below:

1. God exists in time.

2. If God exists in time, He determines the existence of some future entities.

3. Therefore, God determines the existence of some future entities (from 1 & 2).

4. If God determines the existence of some future entities, propositions about those future entities are true.

5. Therefore, some propositions about the future are true (from 3 & 4).

6. A proposition about an entity whose existence is caused by the direct intervention of God is made true by the existence of that entity.

7. If propositions about entities whose existence are caused by the direct intervention of God are true, it entails that those entities exist.

8. Therefore, some future entities exist (from 6 & 7).

9. If any wholly past or merely future entities exist, presentism is false.

10. Therefore, presentism is false.

Presentism is inconsistent with premise (1), the belief that God exists in time. Premise (1) is assumed as true. The defence of the argument will be divided into two sections. Section 3 will defend premises (2) to (5). Section 4 will defend premises (6) to (8). The entities referred to in premises (6) and (7) are defined as 'entities who will exist in the future because of the direct intervention of God in the present moment', this will be abbreviated to E for the rest of this paper.

---

7. *Ibid.*, 53.

3   Premises (2) - (5)

Premise (2) states that if God exists in time, He determines the existence of some future entities. This section will prove premise (2) and then will explain the steps taken to reach premises (4) to (5). The following scenario explains God's effect on the future, which results in premise (2). God can perform an action X in the present moment. If the purpose of action X is the intended outcome Y, it is the case that Y will occur. In other words, if God performs action X with the intention of Y, Y will happen. The outcome could be the existence of an entity. Hence, God could act in the present moment to bring about the existence of an entity in the future. It should be noted that similar arguments have previously been made. For example, it has been argued that if God wills for a future event to happen then that event will happen.[8] I will now explain, in greater detail, how action X leads to the outcome Y, showing that the existence of some future entities is determined by God's actions in the present moment.

If God performs action X with the intention of effect Y, Y will happen:

1∗.   God performs action X.

2∗.   God's sufficient reason for performing action X is the outcome Y.

3∗.   God has the sufficient reason to act in each moment leading up to Y, to ensure that Y will happen.

4∗.   God is omniscient and consequently has the knowledge in each moment leading up to Y, to know how to make Y happen.

5∗.   God is omnipotent and can accordingly directly intervene in each moment leading up to Y, to ensure Y happens.

6∗.   Therefore, it is the case that the Y will occur (from 3∗, 4∗ & 5∗).

Premises (2∗) to (3∗) explain why outcome Y must occur after action X has been performed. Theists maintain that God's actions always have a reason because they would not have faith in a whimsical being. God's sufficient reason for performing action X is Y. Consequently, if Y does not happen then action X would be arbitrary, an unacceptable outcome for theists. For example, if it was argued that after performing X God changed His mind, and brought about an outcome different to Y, it would subsequently have to be admitted that God performed a pointless action. Further actions might be required after X to cause Y, such as stopping future events which could prevent Y. God is unable to directly act in the future from the present moment because

---

8. Thomas Aquinas, *Summa Theologica*, trans. Fathers of the English Dominican Province (N.Y.: Benziger Brothers, 1947).

the future does not exist. Hence, premise (3∗) states that God can act in every moment leading up to Y, to prevent possible events from affecting the existence of Y.

Premises (4∗) and (5∗) both use God's attributes to show how further actions could be carried out to ensure that Y will happen. Premise (4∗) utilizes God's omniscience to demonstrate that in each moment leading up to Y, God knows how to make Y happen. Premise (5∗) refers to the attribute of omnipotence, which is defined as God being all powerful in the present moment. God has the power, in each moment leading up to the outcome, to ensure that the outcome will happen. As a result of premises (2∗) to (5∗), after action X has been performed with the intention of bringing about Y, it is the case that Y will occur.

To summarise, God can perform an action in the present moment which will bring about an outcome in the future. This could be the existence of an entity, thus proving premise (2). If God exists in time, He determines the existence of some future entities, resulting in premise (3), God determines the existence of some future entities. If the existence of some future entities is determined, those entities will exist in a certain way. Propositions which refer to how those entities will exist are true. This leads to premise (4), if God determines the existence of some future entities, propositions about those future entities are true. Premise (5) follows from premise (4), there are some true propositions about the future.

4   Premises (6) - (8)

This section has been divided into two sub-sections. Section 4.1 introduces truthmaker theory which is used in premise (6); a proposition about an entity whose existence is caused by the direct intervention of God is made true by the existence of that entity. It will explain the problem that this causes for presentism. Section 4.2 discusses anticipated objections to premise (6). These objections will deny that propositions about E require the existence of E to be made true. The end of this section explains how premise (6) leads to premise (7), which results in premise (8).

4.1   Premise (6)

Premise (6) refers to truthmaker theory, which states that true propositions require a truthmaker, something in virtue of which they are true.[9] Propositions which refer to a portion of reality can be made true by a truthmaker, the corresponding part of reality. Hence, the truth of a proposition, by the means of a truthmaker, is grounded

---

9. Gonzalo Rodriguez-Pereyra, "Why Truthmakers", in Truthmakers: *The Contemporary Debate*, ed. Julian Dodd (Oxford: Clarendon Press, 2005), 17.

in reality.[10] Propositions which do not have a truthmaker are ungrounded, being true or false regardless of reality. Such propositions are avoided in philosophy because they lack justification and do not refer to fact. Truthmaker theory is an evaluative tool used to rule out "dubious ontologies"[11] which are "unwilling to accept an ontology robust enough to bear the weight of the truths [they feel] free to invoke".[12] If a theory cannot supply truthmakers for propositions which are intuitively true, the theory 'cheats' its ontological commitments.

Truthmaker theory challenges presentism's fundamental claim that the past and future do not exist.[13] Most presentists assert that propositions about the past and future are true. However, they are unable to account for the truth of these propositions because they state that past and future events do not exist. Hence, propositions which are about the past and future lack the existence of a past or future event to make them true. For example, the proposition 'Julius Caesar existed' lacks a truthmaker because Julius Caesar does not exist. The proposition is ungrounded which is an unacceptable conclusion for an advocate of truthmaker theory. It is stated in premise (6) that a proposition about E is made true by the existence of E. Presentism is unable to account for the truth of these propositions because E does not exist in the present moment.

The presentist has three options to avoid the truthmaker problem; reject truthmaker theory, deny that there are true propositions about the past and future, or locate a truthmaker in the present moment. I have assumed that truthmaker theory is true, and it has been proved that there are true propositions about the future. Hence presentists are left with the last option; to locate a truthmaker in the present moment. In the following section, I present two theories which have attempted this and adapt them to object against premise (6). If they are successful, propositions about E would not require the existence of E to be made true.

4.2   Defending Premise (6)

First objection: Ned Markosian has argued that truthmaker theory does not present a problem against presentism.[14] Truthmakers for propositions about the future are comprised of presently existing objects, their current arrangement, and the laws of nature. To begin with, I summarise Markosian's argument. This is followed by its adaptation which will object to premise (6). Markosian stated truthmaking theory as the following:

---

10. Rodriguez-Pereyra, "Why Truthmakers", 21.

11. Theodore Sider, "Against Presentism", in *Four-Dimensionalism: an Ontology of Persistence and Time* (Oxford: Clarendon Press, 2001), 40.

12. Sider, "Against Presentism", 41.

13. Ned Markosian, "The Truth About the Past and the Future," in A*round the Tree: Semantic and Metaphysical Issues Concerning Branching and the Open Future,* ed. Fabrice Correia and Andrea Iacona, vol. 136 (Dordrecht: Springer, 2013), 128. https://link.springer.com/chapter/10.1007/978-94-007-5167-5_8.

14. Markosian, "The Truth About the Past and the Future."

The Truthmaker Principle: For every truth, p, there exist some things $x_1, ..., x_n$, such that p is true in virtue of the existence and arrangement of $x_1, ..., x_n$. [15]

Markosian argued that presently existing objects are one component of a truthmaker. They ground propositions about the future in the present moment because they are made from the same matter as future entities. The two other components needed to make a proposition true are the current arrangement of existing objects and the laws of nature. The current arrangement of existing objects will be determined to exist in a certain way by the laws of nature. This is the extent to which true propositions can be made about their future existence. Notably, Markosian argued that the laws of nature are unable to determine the existence of all future entities. In the present moment, the propositions that refer to these entities are false. Hence, all three of the components described above can make some propositions about the future true. Markosian stated that this would work in the following way:

In 1,000 years there will be human outposts on Mars.

The laws of nature will determine the current arrangement of objects to form human outposts on Mars in 1,000 years. Propositions about human outposts on Mars are grounded in the present moment by presently existing objects which will comprise outposts. The proposition 'In 1,000 years there will be human outposts on Mars' is made true by all three components.[16] Markosian's argument could offer a solution to the truthmaker problem for presentists.

For Markosian's argument to object against premise (6), it must be adapted to supply truthmakers for propositions about E. This could state that a proposition about E is made true by presently existing entities, their arrangement and determined existence. The laws of nature can be changed to God's actions in the present moment because they both play the role of determining the existence of some future entities. As explained in section 3, if God acts in the present moment to bring into existence E in the future, the existence of E is determined. After the initial action, there are true propositions about E. The existence of E is also determined by the current arrangement of existing objects. God's decision to bring E into existence depends upon what exists in the present moment. Hence, God's action in the present moment and the current arrangement of objects determines the truth values of propositions about E. Presently existing objects ground propositions about E in the present moment because they will comprise E. All three components make a proposition about E true in the present moment. Hence, the adaptation of Markosian's argument states that propositions about E can be made true by truthmakers which are located in the present moment.

---

15. *Ibid.*, 130.
16. *Ibid.*, 131.

This objection fails because presently existing objects are unable to ground propositions about E. Markosian argued that propositions about the future can be grounded by presently existing objects because they are made of the same matter as future objects. However, God can change or create matter without limitations because He is omnipotent. Consequently, E could be made from a different combination of matter or from an entirely new matter. In these scenarios, propositions about the future could not be grounded by presently existing matter. For example, God could commit action X in the present moment, the combining of two colours, which in the future will create a new colour. Once the colour is created, it is possible that it would not resemble any existing colours and could not be made from existing colours. If this were the case, there would not be matter in the present moment which could ground propositions about the new colour. A proposition about the colour would require the future colour to act as a truthmaker and ground the proposition. For this to be the case, the colour would need to exist. Hence, this objection fails because it is possible for E to be made of a different matter than presently existing objects. In this case, presently existing objects would be unable to ground propositions about E. Therefore, it has been shown that Markosian's adaptation does not adequately ground propositions about E in the present moment.

The flaw in Markosian's argument demonstrates that propositions about E can only be made true by the existence of E. God's power is limitless in the present moment. In each second leading up to an outcome, He is able to bend and change matter. The new matter could therefore be entirely different to the matter that existed in the previous present moment. Consequently, only the existence of E can be a truthmaker for propositions about E. So far, premise (6) has been shown to be true; a proposition about E can only be made true by the existence of E.

Second objection: Rhoda has argued that God's existence in time can overcome the truthmaker problem for presentism.[17] God's representational mental states of the past can serve as truthmakers to make propositions about the past true. This argument can be summarised: God remembers the past, experiences the present moment and anticipates the future. As God is omniscient, He accurately remembers His experiences of the past. These memories are concrete representational mental states which can serve as truthmakers for propositions about past events.[18] For example, God remembers how Julius Caesar existed. Accordingly, His memories of how Caesar existed are representational mental states formed from His experiences of Caesar. God's representational mental state of Julius Caesar being assassinated can serve as a truthmaker for the proposition that 'Julius Caesar was assassinated'. Hence, Rhoda has argued that propositions about the past can be made true by God's representational mental states.

Rhoda's argument can be adapted to supply truthmakers for propositions about E.

---

17. Rhoda, "Presentism, Truthmakers, And God".
18. Rhoda, "Presentism, Truthmakers, And God", 54.

Propositions about the future can be made true by God's representational mental states of future events. As a result of the attribute of omniscience, God could form representational mental states of what future events would be like based on His experience of viewing the past. As God creates E, He knows how E will exist and accordingly is able to form mental states of E. These mental states could act as truthmakers to make propositions about E true.

The problem with this adaptation is that God's mental states, which do not represent how reality will be, could make false propositions true. God's mental states, whether or not they represent future reality, exist in the same capacity. So far, no reason has been given which could explain why God's mental states which represent the future are able to make propositions true and not so the mental states which do not represent the future. For this adaptation to successfully object against premise (6), only God's mental states which represent how reality will be should be able to make propositions true.

It could be argued that only God's mental states which God knows represent the future can function as truthmakers. If God knows that some of His mental states are inaccurate about the future, He knows that they do not represent any facts about the future. God would be unable to know this by knowing one fact about the future. Rather, God would have knowledge of the relevant facts about the future and could use these facts to infer or deduce which future events are not going to happen. These mental states could be categorised as non-factual states, meaning that they do not represent any facts about the future which are known by God. If God knows that some of His mental states accurately represent the future, He knows the future events which they represent are going to happen. These could be categorised as factual mental states, meaning they represent facts known by God about the future. It could be argued that God's mental states which do not represent the future are unable to function as truthmakers, such that God's non-factual mental states do not represent any facts about the future which are known by God, and consequently are unable to ground propositions about the future in facts. If the future is partly determined, God may not know if some of His mental states represent the future because He would not know if the future events that they represent are going to happen. These would be non-factual mental states because they do not represent any future facts which are known by God and therefore would be inadequate truthmakers. Hence, only God's mental states which He knows represent the future can act as truthmakers. God knows mental states about E are accurate because He knows E will exist. Mental states of E are categorised as factual states. Hence, God's mental states of E could make propositions about E true.

This objection fails because God's foreknowledge does not affect the contents of His representational mental states which makes propositions true. According to the above argument, God's foreknowledge can be used to assess whether His mental states are factual or non-factual. On the other hand, propositions are made true by the contents of God's representational mental states. God knowing that a mental state misrepres-

ents or represents the future does not affect the contents of His mental states. Consider God's mental state of a red flower existing. As God knows a red flower will exist, this mental state is factual. Propositions about a red flower existing can be made true by God's mental state of a red flower. God could also have a mental state of a yellow flower. God could know that there will not be a yellow flower. According to the above argument, propositions about a yellow flower could not be made true by God's mental state of a yellow flower because they do not represent any facts. However, God's knowledge that the first mental state is factual and the second is non-factual does not affect the contents of the mental states in either example. Regardless of this division, the mental states exist in the same capacity. If the first mental state is able to function as a truthmaker, the second mental state should also be a truthmaker. Hence, God's mental states which misrepresent the future should be able to make propositions true. As stated above in the objection, God's mental states which misrepresent the future should not be able to make propositions true because those propositions would be false. Hence, the adaptation of Rhoda's theory fails because it allows for false propositions to be made true by God's mental states.

To summarise section 4.2, the objections above fail because they do not demonstrate that propositions about E can be made true by anything other than the existence of E. The adaptation of Markosian's theory is unsuccessful because presently existing objects are unable to ground propositions about E. The adaptation of Rhoda's theory fails because it allows for false propositions to be made true. Moreover, Markosian's objection shows that only the existence of E can serve as a truthmaker for propositions about E. Hence, premise (6) is proved to be true; a proposition about E can only be made true by the existence of E.

Premise (7) states that if propositions about E are true, it entails that E exists. Section 3 has proved that propositions about E are true. It has been shown that propositions about E can only be made true by the existence of E. It follows that E exists. Consequently, premise (8) is true; some future entities exist. This is inconsistent with presentism as it states only presently existing entities exist. Therefore, presentism is inconsistent with the belief that God exists in time.

5   Conclusion

This paper has shown that the belief that God exists in time is inconsistent with presentism. In section 3, premise (2) was proved by the argument that if God performs action X with the intention of effect Y, Y will happen. Premises (4) and (5) followed which led to the conclusion, if God exists in time, there are true propositions about the future. This should be relatively uncontroversial as Aquinas has produced a similar and well-

established argument.[19] Furthermore, it has been clearly explained how this argument works within the framework of presentism. This led to section 4 which explained and defended premises (6) to (8). Section 4.1 introduced truthmaker theory which is used in premise (6). I have assumed that truthmaker theory is true as it is generally accepted by both presentists and non-presentists. To overcome the truthmaker problem, presentists must locate truthmakers for propositions about the future in the present moment. In section 4.2, I explained how presentists might attempt to do this for propositions about E. I have dedicated a large amount of this paper to explaining and rejecting these objections because I expect that they will be used as the main argument against my conclusion. I conclude in section 4.2 that propositions about E can only be made true by the existence of E. The last part of section 4 explained premise (7); if propositions about entities whose existence is caused by the direct intervention of God are true, it entails that those entities exist. This resulted in premise (8); some future entities exist. Hence, if God exists in time then future entities exist. Presentism states that only presently existing objects exist. Therefore, this paper has shown that presentism is inconsistent with the belief that God exists in time. In other words, God cannot only exist in the present moment.

References

Aquinas, Thomas. *Summa Theologica*. Translated by Fathers of the English Dominican Province. N.Y.: Benziger Bros., 1947.

Craig, William Lane. *Time and Eternity: Exploring God's Relationship to Time*. Wheaton: Crossway Books, 2001.

Markosian, Ned. "A Defence of Presentism". Essay. In *Oxford Studies in Metaphysics; Volume 1*, edited by Dean W. Zimmerman, 47–82. Oxford : Oxford University Press, 2004.

Markosian, Ned. "The Truth About the Past and the Future." Essay. In *Around the Tree: Semantic and Metaphysical Issues Concerning Branching and the Open Future* 136, edited by Fabrice Correia and Andrea Iacona, 136: 127–41. Dordrecht: Springer, 2013. https://link.springer.com/chapter/10.1007/978-94-007-5167-5_8.

Rhoda, Alan R. "Presentism, Truthmakers, And God." *Pacific Philosophical Quarterly* 90, no. 1 (2009): 41–62. https://doi.org/10.1111/j.1468-0114.2009.01328.x.

Rodriguez-Pereyra, Gonzalo. "Why Truthmakers." Essay. In *Truthmakers: The Contemporary Debate*, by Beebee, Helen, and Julian Dodd, eds., edited by Helen Beebee, and Julian Dodd, 17–32. Oxford: Oxford University Press, 2005.

---

19. Aquinas, *Summa Theologica*.

Sider, Theodore. "Against Presentism." Essay. In *Four-Dimensionalism: an Ontology of Persistence and Time*, 40–41. Oxford: Clarendon Press, 2001.

A Resentment of Disappointment for the Politics of Resentment

LOCHLANN ATACK*
University of Edinburgh

Abstract   Prejudicial beliefs are often associated with ignorance. Indeed, common views of prejudice hold that it is precisely in their ignorance that those with prejudicial beliefs perpetrate a wrong to the victim of the prejudicial belief. This view of prejudice neatly accounts for the prejudices of epistemically culpable epistemic agents. But what about cases where a prejudice is held by an epistemically exculpable epistemic agent? This paper presents an example of a deeply prejudiced belief about Indigenous Peoples, taken from a recent ethnography of rural Wisconsin, and argues that it is epistemically exculpable. If it is indeed epistemically exculpable, then we need to look beyond the individual when directing our blame for the prejudicial belief; we can only be disappointed in the circumstances that enable an epistemic agent to be epistemically exculpable for expressing such a belief.

1   Introduction

This paper attempts to give an epistemological account of the phenomenon identified as "the politics of resentment" by the sociologist Katherine Cramer, which "arises from the way social identities, the emotion of resentment, and economic insecurity interact".[1] Specifically, the paper argues that we cannot epistemically blame certain individuals within the politics of resentment for their extremely prejudiced, morally blameworthy beliefs. In section 2, I present the central case under discussion in the paper as

---

*Lochlann Atack has recently completed his undergraduate degree in philosophy at the University of Edinburgh. Particularly interested in epistemology, political theory, and critical pedagogy, his entry onto Teach First's Training Programme in London has been deferred until September 2021 due to the COVID-19 pandemic. In the meantime, he looks forward to spending a civically engaged year in his native city of Dublin.

1. Katherine J. Cramer, The Politics of Resentment: Rural Consciousness in Wisconsin and the Rise of Scott Walker (Chicago: University of Chicago Press, 2016), 9.

it appears in Cramer's ethnography of political consciousness in rural Wisconsin: a rural Wisconsinite's deeply prejudiced belief about Indigenous Peoples.[2] [3] In section 3, I motivate Endre Begby's challenge to a common view of prejudice, which shows that prejudiced belief can be epistemically justified, and his notion of *highly non-ideal epistemic contexts* as sufficient circumstances for epistemic exculpation of prejudiced beliefs. In section 4, I show how we can extend Miranda Fricker's condition for epistemic exculpation of testimonial injustice, historical-cultural distance from the present moral discourse, to the central case in this paper. Ultimately, this paper suggests a counterintuitive, perhaps troubling conclusion about prejudicial beliefs produced in the "politics of resentment"; that, in Fricker's terms, we cannot extend a "resentment of blame" to those with prejudicial beliefs, but only a "resentment of disappointment".[4]

2   The Politics of Resentment: The Central Case

In the period between May 2007 and November 2012, the sociologist Katherine Cramer visited 27 different community groups, mostly in rural communities, across Wisconsin, to listen to their conversations.[5] In doing so, Cramer's "main motivation was not to get at how well people make sense of politics, but to get at how they do so".[6] Cramer's findings encompassed a broad spectrum of issues, from views on economics, to healthcare, to the University of Wisconsin-Madison.[7] Here, I focus on the findings with distinctively high moral stakes: the prejudicial beliefs that racially dominant groups have on racially dominated groups. For the purposes of the paper, I take it for granted that

---

2. I want to flag in advance that I will be quoting directly this lengthy, dehumanising prejudicial belief, and to emphasise that discussion of the belief will occur throughout the paper. The quote begins on the next page. My thanks to Dr. Aidan McGlynn for guidance on flagging this aspect of the paper, and for prompting my inclusion of the following footnote.

3. It is important to acknowledge at the outset of this paper that the appropriate collective term of reference for so-called "Indigenous Peoples" is controversial and contested. The prejudiced belief under discussion, since it does not specify a specific tribal identity, all but forces my hand to use this collective term which, although less harmful than some other terms, is nonetheless directly connected to the centuries-long, ongoing practices of settler-colonialism perpetrated against Indigenous Peoples. For an excellent introduction to the importance of this issue, see Michael Yellow Bird, "What We Want to Be Called: Indigenous Peoples' Perspectives on Racial and Ethnic Identity Labels," American Indian Quarterly 23, no. 2 (1999): p. 1, https://doi.org/10.2307/1185964. For work on the role of misrepresentation in the marginalization of Indigenous Peoples in contemporary higher education, see Adrienne Keene, e.g. "Representations Matter: Serving Native Students in Higher Education," Jcscore 1, no. 1 (June 2018): pp. 101-111, https://doi.org/10.15763/issn.2642-2387.2015.1.1.101-111. I am very grateful to members of the "Decolonise UWC" Facebook group for guidance and literature recommendations on this issue.

4. Miranda Fricker, Epistemic Injustice: Power and the Ethics of Knowing (Oxford; New York: Oxford University Press, 2007), 104.

5. Cramer, The Politics of Resentment, 29, 35.

6. *Ibid.*, 20.

7. *Ibid.*, 37, 46, 210.

white Americans constitute the racially dominant group, and non-white Americans constitute racially dominated groups. Thus, I take as the central case for the paper the views of "Ron", a white "man in the group of loggers in northwest Wisconsin", speaking about "American Indians" who live on the local reservation.[8] [9] Here is Ron's view of the Indigenous Peoples living in the local reservation:

> And, there, there's too many programs down there for a bunch of people, you know to have it for them to want to go to work. You know? They got the casino down there shoving our money through 'em, they got the federal government shoving our money through 'em, and they wonder where they got drunken alcohol problems, they got nothing to do all day besides sitting around and do what they want to do. And they keep giving 'em money to do it, well how do you expect to get anything out of anybody? There's an old saying: A hungry dog hunts harder. Hey, you keep feeding a dog or a cat, they're not gonna hunt, they're not gonna look for food, they're gonna lay around and get fat. [10]

Herein, this case will be referred to as 'RON'. Now, perhaps somewhat counterintuitively, when Cramer listened to Ron express this racist view—complete with a likening of Indigenous Peoples living in reservations to lazy, fat domestic animals—she didn't take it to be an expression of "just plain racism". Rather, for Cramer, taken in context, Ron's racism is baked into "a much more complex process of sense making" that constitutes how he "make[s] sense of public affairs".[11] Or put more bluntly, as Cramer sees it, Ron isn't being racist for the sake of it, or simply because he's a bigot, but because his understanding of society requires it. Cramer understands this kind of racism as a consequence of the sense of identity unique to rural Wisconsinites, which she calls "rural consciousness", and the "broader structural forces" they are subject to, which together amount to a "politics of resentment".[12] Therefore, to sufficiently understand cases like RON requires a consideration of the role that Ron's location in a politics of resentment plays. Cramer's diagnosis of this kind of prejudice, I will now argue, finds support in the recent work of Endre Begby on epistemic responsibility.

---

8. *Ibid.*, 88.

9. As mentioned in footnote 3., the nature of Ron's prejudiced belief makes it difficult to specify with certainty what specific reservation he is referring to. There are twelve federally recognised First Nations within the state of Wisconsin, and most of northern Wisconsin lies on Ojibwe land—however there are several reservations within just north-west Wisconsin so it is impossible to know with certainty which "local" reservation Ron has in mind. For a map of the reservations in Wisconsin, see "Current Tribal Lands Map and Native Nations Facts," Wisconsin First Nations, March 31, 2020, https://wisconsinfirstnations.org/current-tribal-lands-map-native-nations-facts/.

10. Cramer, The Politics of Resentment, 88.

11. *Ibid.*, 18, 89.

12. *Ibid.*, 24.

3   Begby's Account of Epistemic Culpability

Begby's account of epistemic responsibility supports Cramer's diagnosis of RON by:
(a) problematising a "common view" of prejudice and (b) developing the notion of
highly non-ideal epistemic contexts.

3.1   Problematising a "common view" on prejudice

Begby defines a *prejudice* as "a negatively charged, materially false stereotype target-
ing some social group and, derivatively, the individuals that comprise this group".[13]
Working under this definition, Begby challenges a "common view" of prejudice ad-
vocated for by both Nomy Arpaly and Miranda Fricker.[14] Roughly, this common view
holds that prejudices "always arise from some specifiable mishandling of evidence"
such that they are deemed epistemically culpable and irrational.[15] Begby challenges
this view on the basis that it "takes for granted a characterization of prejudices in terms
of universal generalizations" and that this assumption is incorrect "at least for the ma-
jority of common and pernicious prejudices".[16] This assumption, Begby argues, elides
the distinction between "generic judgements" and "universal generalizations" when
identifying stereotype judgements. i.e. it assumes stereotypes make universal gener-
alizations about groups, rather than fallible generalizations.[17] Thus "all immigrants
are illegal aliens" is a universal generalization, but "immigrants are illegal aliens" is a
generic judgement. This, albeit subtle, distinction is crucial because, as Begby points
out, universal generalizations are far more sensitive to negative instances than generic
judgements.[18] A single negative instance is enough to comprehensively undermine a
universal generalization: if I come across a "legal" immigrant, then that is sufficient
evidence to undermine, via falsification, the universal generalization "all immigrants
are illegal aliens", since that proposition is incompatible with the proposition "this im-
migrant is legal." However, the generic judgement "immigrants are illegal aliens" is
not falsifiable in the same way, even in the face of the same evidence. The proposition
"immigrants are illegal aliens" is compatible with "this immigrant is legal", since we
can understand the generic judgement to express a tendency, rather than a necessary
condition, of a group. Thus the generic judgement that immigrants tend to be illegal
aliens can accommodate a couple of "legal" immigrants here and there in a way that a
universal generalization cannot. The upshot of this distinction is that while it is obvi-
ously an "epistemic mistake" to sustain a falsified universal generalization, it is much

---

13. Endre Begby, "The Epistemology of Prejudice," Thought: A Journal of Philosophy 2, no. 2 (2013):
pp. 90-99, https://doi.org/10.1002/tht3.71, 90.
14. Begby, "The Epistemology of Prejudice", 90.
15. *Ibid.*
16. *Ibid.*, 91.
17. *Ibid.*
18. *Ibid.*, 94.

less clear that to sustain a generic judgement in the face of a counterexample constitutes such a mistake.[19] In this way, prejudices, "once they are internalized... can quite reasonably come to control the assessment and interpretation of new evidence".[20] This aspect of "epistemic insidiousness" is not captured by the common view endorsed by Fricker and Arpaly.[21]

For illustration, Begby refers to Arpaly's hypothetical case of Solomon, "a boy who lives in a small, isolated farming community in a poor country" who believes that women are significantly less intellectually able than men.[22] The "common view" holds that while in this context, Solomon is not epistemically culpable for this belief, and so cannot be said to be prejudiced, for lack of evidence to suggest otherwise.[23] However, the common view holds that once Solomon encounters countervailing evidence to this belief (for example, at university where there are countless women who are at least as intellectually capable as men) then he becomes epistemically culpable for his belief.[24] If Solomon's belief relies on a universal generalization that "all women are intellectually inferior to men", then it seems the prediction of the common view holds, since he has encountered countless exceptions to this generalization at university. However, if his belief relies on a generic judgement about women's intellectual abilities compared to men's, then it can be rationally revised to something like: "only women who study at university can be the intellectual equals or superiors of men." Indeed, this actually seems like the more rational response to his evidence. For Solomon would be revising his belief about women proportionate to the countervailing evidence: he has encountered evidence to suggest that "women at university are intellectually capable", but has not encountered evidence to suggest that "women not at university are intellectually capable." To form the latter belief, then, would be to form a belief without evidence. The upshot here is that Solomon seems epistemically justified, rather than culpable, for maintaining the generic judgement that women tend to be less intellectually capable than men. But this, clearly, is still a false stereotype that wrongs women. Thus, Begby concludes that we need to "move beyond" the common view "to determine the exact nature of Solomon's prejudice".[25] More precisely, if we are to understand "how prejudices are capable of absorbing or subsuming such contrary evidence, we must recognize that they paint a more complex picture of the world than the common view assumes".[26]

Begby's treatment of Solomon is remarkably consonant with Cramer's judgement of RON. As Cramer notes, when we fully contextualize RON we see that Ron is making

---

19. *Ibid.*, 92.
20. *Ibid.*, 91.
21. *Ibid.*, 97.
22. *Ibid.*, 91.
23. *Ibid.*
24. Nomy Arpaly, Unprincipled Virtue: an Inquiry into Moral Agency (Oxford: Oxford University Press, 2011), 104.
25. Begby, "The Epistemology of Prejudice", 92.
26. *Ibid.*, 97.

a comparison between an "American Indian" he works with, and those that live on reservations:

> RON: Yeah. You know. Well like him that just left, that was here before to get coffee?
>
> KJC: Yeah.
>
> RON: He's an American Indian. [One sentence deleted for confidentiality.]
>
> KJC: Oh really?
>
> RON: He's a good guy.
>
> KJC: Yeah.
>
> RON: Works hard. Yeah.
>
> KJC: Well sure.
>
> RON: But he won't live on the reservation where they get all that free housing and stuff, he's self-supporting, you know?
>
> KJC: Yeah.
>
> RON: And, there, there's too many programs down there for a bunch of people, you know to have it for them to want to go to work... they're gonna lay around and get fat." [27]

When we put Ron's prejudiced tirade in context, it appears to be a paradigmatic example of a stereotype premised on a generic judgement, and not a general universalsation. RON does not express that "all Indigenous Peoples are lazy alcoholics", but, rather, that "Indigenous Peoples who live on reservations are lazy alcoholics". As we can see, this stereotype permits him to rationally hold that some Indigenous Peoples can be "good guy[s]" with a good work ethic, while others have a work ethic that, Ron believes, permit dehumanising comparisons with domestic animals. It doesn't seem unreasonable to suppose that perhaps, at some point, Ron didn't believe Indigenous Peoples to be capable of having a good work ethic *at all*, but, upon working with one of them, revised his belief accordingly in the face of that evidence. i.e. just as Solomon, while at university encountered evidence that the women there were intelligent, Ron, while at work encountered evidence that the Indigenous Person there had a good work ethic. Thus, just as with Solomon, to expect Ron to revise his belief further than his evidence permits is to expect him to act irrationally. As counterintuitive and troubling as it may seem, on this analysis Ron is not epistemically culpable. Hence, Begby's treatment of the Solomon example gives Cramer's claim that "To call [RON]

---

27. Cramer, The Politics of Resentment, 88.

just plain racism misses the complexity of the sentiments involved here" an epistemic basis.[28] To call Ron's prejudice "just plain racism" does not recognise the epistemic insidiousness of his belief, and thus provides an oversimplified diagnosis of the wrong that his prejudice perpetrates.

3.2   Epistemic agents in "highly non-ideal epistemic contexts"

Elsewhere, Begby develops the notion of "highly non-ideal [epistemic] contexts" (HNIEC).[29] Paradigmatic examples of HNIEC are those where people "grow up in deeply prejudicial social settings, with no rational access to contrary evidence".[30] In such cases, epistemic agents' prejudiced beliefs are simply "the predictable consequences of their limited epistemic opportunities".[31] In this way, they are "victims" with the "peculiarly bad luck of growing up in a severely constrained socio-epistemic environment".[32] Thus on Begby's account of doxastic responsibility, being raised in a HNIEC is a sufficient condition for epistemic exculpation of a prejudiced belief.

Now, recall that Cramer's entire project is premised on the motivation not to evaluate, but simply to understand, how rural Wisconsites "make sense of politics".[33] While not made explicit, this approach seems to presuppose, on Cramer's part, an understanding that these people are not operating in ideal epistemic contexts. Specifically, Cramer seems to identify two distinct ways in which rural Wisconsinites like Ron are subject to HNIECs. Indeed, I will now show that not only is one's rural consciousness "materially" reinforced through the everyday lived experiences of those with a similar share of political power, like one's family or colleagues, but it can also be "symbolically" exploited by those with a higher share of political power, like politicians.

3.2.1   Reinforcement of prejudice

Essential to rural consciousness is the sense that rural Wisconsinites are victims of distributive injustice, and that "rural folks are fundamentally different from urbanites in terms of lifestyles, values, and work ethic".[34] In a 2011 state-wide opinion poll, Cramer found that "69 percent of rural respondents felt rural areas received much less or somewhat less than their fair share" of government resources.[35] Cramer shows that this claim is unequivocally false: "the evidence certainly does not support the notion

---

28. *Ibid.*
29. Endre Begby, "Doxastic Morality," Philosophical Topics 46, no. 1 (2018): pp. 155-172, https://doi.org/10.5840/philtopics20184619, 166.
30. Begby, "Doxastic Morality", 168.
31. *Ibid.*
32. *Ibid.*
33. Cramer, The Politics of Resentment, 24.
34. 35.*Ibid.*, 5.
35. *Ibid.*, 105.

that urban counties receive far more than their share of tax dollars per resident" —in fact "rural counties in the aggregate pay somewhat less in taxes per person and receive approximately similar amounts of money in return".[36]

However, rather than concluding from this that rural consciousness is fundamentally an irrational, ignorant social identity, Cramer provides evidence that such communities, while not victims of distributive injustice, face considerable pressures. First, rural counties "tend to experience greater levels of poverty, lower wages, and modestly higher levels of unemployment".[37] Second, rural communities have been "experiencing a long, slow death for decades", and fighting a "losing battle economically" —local businesses, the lifeblood and pride of rural communities, cannot compete with franchises and the rapidly expanding global economy.[38] Third, "providing broadband service is a more daunting task in a sparsely populated community than it is in a dense urban one" due to "economies of scale", and "[e]ducation levels tend to be lower" in rural communities.[39]

Taken individually, these facts will present their own obstacles to epistemic practices—limited broadband will hinder access to online sources of information, for instance. When taken together, these facts constitute an extremely constrained socio-epistemic environment and thus will shape one's epistemic character. Moreover, if one is interacting with people subject to the same constraints, the beliefs emerging from such an environment will be mutually reinforcing. Thus, while it may be true that rural consciousness, as expressed by those who share it, is premised on fiction, that is only part of the epistemic story here. If we are misguided when we judge Solomon's epistemic character without having considered the epistemic context he was raised in, the same applies to our judgement of RON. Ron, as a rural Wisconsinite, is statistically more likely to have experienced greater levels of poverty, lower wages, less employment, lower levels of education, and been raised in a dying community, and is surrounded by people who are not likely to challenge his beliefs. If this doesn't qualify as a HNIEC, it's unclear what would. And if being raised in a HNIEC is sufficient for epistemic exculpation, then on this basis we ought to epistemically exculpate RON.

3.2.2   Exploitation of Prejudice

Rural consciousness will not only be shaped by severe material constraints and reinforced by interactions with individuals who share that social identity, like colleagues and family members. RON's socio-epistemic context is also one of "a broader politics in which tapping into resentment is an effective political strategy".[40] Although rural

---

36. *Ibid.*, 91.
37. *Ibid.*, 94.
38. *Ibid.*
39. *Ibid.*, 99.
40. *Ibid.*, 21.

consciousness is "not something that any one politician instilled in people overnight—or even over a few months", political actors "mobilize support" by "tapping into" its latent resentment when it may be politically advantageous.[41] According to Cramer, politicians like Scott Walker "made use of the desire for people to make sense of their world, to figure out who is to blame and identify boundaries that clearly show that those who are to blame are not one of us" to win gubernatorial elections.[42] In this case, rather than challenging rural consciousness, Walker encouraged it for his political gain.

Thus, there is a mutually reinforcing *vertical* relationship between those in power and those who give power. The former have motivation to rather than challenge, exploit the prejudices of rural consciousness, while the latter have motivation to vote for those who they feel represented by: those political actors who do not challenge their social identity and way of life. Cramer characterises this as "a political culture in which political divides are rooted in our most basic understandings of ourselves, infuse our everyday relationships, and are used for electoral advantage by our political leaders".[43] In this context of a politics of resentment, those with the power—both symbolic and material—to challenge prejudicial beliefs are motivated to instead encourage those beliefs.

Thus we can see that the politics of resentment constitutes a HNIEC in two distinct ways. First, the basic material conditions of Ron's life will be highly non-epistemically ideal, and he is not likely to regularly encounter individuals who will challenge his beliefs. Second, Ron's prejudicial beliefs are more likely to be encouraged, rather than challenged, by the political actors he feels represented by. Thus, insofar as RON is a consequence of the politics of resentment, RON is a consequence of a HNIEC, therefore epistemically exculpating RON.

4   Epistemic Injustice in a Politics of Resentment?

Having shown RON to be epistemically exculpable on the basis that the expressed prejudice can be understood as rational, and that Ron is subject to a HNIEC, I now proceed to discuss how cases like RON affect a prominent account of epistemic injustice. Miranda Fricker, in her foundational work on epistemic injustice, appears to only treat historical cases of exculpatory epistemic contexts, and not contemporaneous cases like RON. Hence, as it stands RON is not epistemically culpable according to Fricker's account. I now present and attempt to resolve this *prima facie* problem, ultimately showing RON to be compatible with Fricker's account.

---

41. *Ibid.*, 19, 213, 221.
42. *Ibid.*, 207.
43. *Ibid.*, 211.

4.1   Resentment of Disappointment in mid-20th Century USA

Consider the following characterisation of one of Fricker's central cases of "epistemic injustice" (which we can refer to as 'USA1950s'):

USA1950s: Herbert is a well-to-do middle aged man raised in the United States of America in the first half of the 20th century. At some point in the 1950s, his daughter-in-law, Marge gives testimony to Herbert about the whereabouts of her husband, Herbert's son. To this, Herbert dismissively responds "Marge, there's female intuition, and then there are facts".[44]

Fricker holds that Herbert's treatment of Marge in this case constitutes an epistemic injustice (EIJ). For Fricker, an EIJ is perpetrated when someone is harmed by someone else "specifically in their capacity as a knower".[45] The specific kind of EIJ perpetrated here is a testimonial injustice, which occurs when someone is afforded an identity-prejudicial credibility deficit (IPCD).[46] An IPCD consists in a speaker being given less credibility than they otherwise would have due to their hearer having an identity prejudice against them.[47] In this way, the speaker is harmed specifically in their capacity as a knower because their attempt to communicate knowledge is undermined by their hearer's prejudice against them. Fricker later characterises the perpetration of a testimonial injustice as a failure to exercise the virtue of testimonial justice.[48] Such a virtuous hearer "neutralizes the impact of prejudice in her credibility judgements" which requires a "distinctly reflexive critical social awareness" either "actively" or "spontaneously".[49]

But while Fricker identifies USA1950s as a paradigmatic case of epistemic injustice, she exculpates Herbert since he is in a cultural-historical "setting in which there is little critical awareness of the construction of gender".[50] Fricker takes this lack of critical awareness to result in a lack of "conceptual resources" available to individuals in that context.[51] In the USA in the 1950s, then, where there was such a lack of critical awareness of the construction of gender, on Fricker's view we can say that the impoverished conceptual resources available to that society were conducive to perpetrating identity prejudice against women, rather than cultivating the virtue of testimonial justice with respect to women. This is borne out when we consider our intuitive response to the following case:

USA2019: Herbert is a well-to-do middle aged man raised in the United States of America towards the end of the 20th century. At some point in 2019, his daughter-in-

---

44. Adapted from Fricker, Epistemic Injustice, 9.
45. Fricker, Epistemic Injustice, 1.
46. *Ibid.*, 28.
47. *Ibid.*, 27-28.
48. *Ibid.*, 86.
49. *Ibid.*, 91-92.
50. *Ibid.*, 99.
51. *Ibid.*, 103.

law, Marge gives testimony to Herbert about the whereabouts of her husband, Herbert's son. To this, Herbert dismissively responds "Marge, there's female intuition, and then there are facts."

In USA2019, it seems clear that Herbert is culpable for his IPCD against Marge. In 2019, rather than a lack, there is an abundance of conceptual resources available to individuals for them to develop critical awareness of the construction of gender. For instance, feminist scholarship has since become well-established. Given this, we have grounds for expecting more from the Herbert in USA2019 compared to the Herbert in USA1950s, since, for example, in the former there exist established concepts such as "gender performativity", whereas in the latter there do not. Fricker articulates this difference in Herbert's blameworthiness by distinguishing between "routine" and "exceptional" discursive "moves in a moral discourse".[52] Exercising the virtue of testimonial justice in USA1950s would require an "exceptional" discursive move given the moral discourse of that context, since the "routine" move would have been to perpetrate IPCD against women. Likewise, exercising the virtue of testimonial justice in USA2019 would only require a "routine" move given the moral discourse of today. Fricker holds that in cases where agents are at fault for not making an exceptional move, they are not culpable for failing to make that move since they were not "in a position to know better".[53] [54] Thus, it is only in cases where agents fail to make routine moves that they can be held culpable for the failure to make it, since it is only in these cases that they are in a position to know better. In cases of culpable ignorance, our intuitive "resentment of blame" towards such prejudiced individuals is justified i.e. our resentment of their prejudice is grounded in a legitimate levelling of blame towards them.[55] However, in cases of exculpable ignorance, we are only justified in a "resentment of disappointment".[56] Here, our justification for resentment of prejudiced individuals only extends as far as our disappointment in their failure to make the exceptional move, which is no justification for blame towards them.

This distinction seems to make the right predictions: if I'm a doctor I would be culpable for not knowing how to treat a common, extant virus, but I can hardly be deemed culpable for not knowing how to treat a virus that has not yet been discovered. In the former case I would have failed a routine move for doctors—knowing how to treat

---

52. *Ibid.*, 104.

53. *Ibid.*, 100.

54. There are, of course, many other accounts dealing with blameworthiness in potentially exculpating contexts besides Fricker's. For some of these recent accounts, see e.g. Nomy Arpaly and Timothy Schroeder, "Praise, Blame and the Whole Self," Philosophical Studies 93, no. 2 (1999): pp. 161-188, https://doi.org/10.1023/a:1004222928272., Lisa Bortolotti and Kengo Miyazono, "The Ethics of Delusional Belief," Erkenntnis 81, no. 2 (2015): pp. 275-296, https://doi.org/10.1007/s10670-015-9739-9., Sanford C. Goldberg, "Should Have Known," Synthese 194, no. 8 (2015): pp. 2863-2894, https://doi.org/10.1007/s11229-015-0662-z., and Elinor Mason, "Moral Ignorance and Blameworthiness," Philosophical Studies 172, no. 11 (2015): pp. 3037-3057, https://doi.org/10.1007/s11098-015-0456-7.

55. *Ibid.*, 104.

56. *Ibid.*

common viruses—but in the latter case I have failed to make an exceptional move—discovering a new common virus. Thus, by appealing to the "historical contingency" of Herbert in USA1950s, Fricker justifies his exculpation.[57] However, Fricker's discussion of such exculpatory contexts focusses exclusively on cultures constrained by "historical distance" from the present moral discourse.[58]

4.2   Resentment of Disappointment in a Politics of Resentment

With reference to the foregoing analysis of RON, I now propose that Fricker's distinction between routine/exceptional discursive moves also holds for cultures contemporaneous to the present moral discourse, and thus exculpates the prejudice in RON. But before doing so, I briefly anticipate an objection to this approach.

One might object that Fricker's endorsement of the "common view" of prejudice suggests that her account of epistemic exculpation is incompatible with cases like RON that undermine the common view, i.e., how can I now be endorsing Fricker's view on exculpating certain prejudiced subjects, when I earlier rejected her view as to what constitutes being a prejudiced subject? The response to this objection is simply to bite the bullet and maintain that Fricker's initial verdict about Solomon seems mistaken. Insofar as the common view has been undermined as a plausible definition of culpable prejudice, what it predicts is inconsequential for further analyses of culpability about prejudice. Moreover, Fricker's account of epistemic exculpation is not developed in terms of the nature of the prejudice, but the context in which the prejudice is perpetrated. Therefore, nothing of substance in Fricker's account of exculpation is contingent on the specific nature of the prejudice, and so we can both reject her endorsement of the common view and endorse her notion of exculpatory contexts. Thus, we only need to focus on whether Fricker's distinction between routine/exceptional moves in a moral discourse can apply to non-temporal distance from the present moral discourse. If it can, then historical distance is not a necessary condition for the relevant kind of inaccessibility to be in play on Fricker's view, which clears the ground for cases like RON to be compatible with Fricker's conditions for exculpability.

Indeed, historical distance does not seem necessary for the relevant kind of inaccessibility to be in play. What is relevant to the distance from the present moral discourse in Fricker's analysis is not its metaphysical inaccessibility—that such concepts literally did not exist in USA1950s—but that they are functionally inaccessible. In other words, it is not the metaphysical status *simpliciter*—whether it exists or not—of a concept that is necessary for some agent to be able to access it. It is the metaphysical status of the concept relative to the agent that is necessary for the agent to be able to

---

57. *Ibid.*, 100.

58. But for a passing suggestion about the possibility of "cultural distance" see Miranda Fricker, "The Relativism of Blame and Williams's Relativism of Distance," Aristotelian Society Supplementary Volume 84, no. 1 (January 2010): pp. 151-177, https://doi.org/10.1111/j.1467-8349.2010.00190.x, 167.

access it. Although the non-existence of a concept is a sure way to guarantee its inaccessibility to some agent—since it will be inaccessible to all agents—it is not necessary.

A concept might exist but be inaccessible to an agent all the same, rendering it functionally inaccessible. We can all too easily imagine a situation where a 'Big Pharma' company has discovered a new common virus at a time t, but kept it under lock and key from the rest of the medical community because it is in their financial interest. But although the virus has been discovered at t, we could hardly hold the doctors not privy to its existence at t culpable for not knowing how to treat the virus after t. It is irrelevant to our judgement of the doctors that at t^{-1} the treatment was metaphysically inaccessible and at t^{+1} it was metaphysically accessible. What is relevant is that at t^{-1} it was inaccessible and t onwards it remained inaccessible to all but the Big Pharma company. Hence, to alter our judgement of the ignorant doctors from before t and after t would be to hold them accountable for the actions of the Big Pharma company—for an action they were completely unaware of. This seems like a very wrongheaded conceptualization of epistemic responsibility. Thus, if mere metaphysical accessibility is not sufficient for moral responsibility, then historical distance is not necessarily the only kind of distance required for epistemic exculpation. That is, the doctors in the above example are distant in a non-historical sense from the knowledge of the virus treatment, a sense we might perhaps call "organisational distance". This organisational distance seems sufficient to exculpate them for their ignorance in this case.[59]

I submit that a similar kind of distance applies to Ron's ignorance, which leads him to RON. Ron's lack of conceptual resources for avoiding prejudice against Indigenous Peoples on the local reservation is not metaphysical – the conceptual resources to not stereotype Indigenous Peoples certainly do exist. But Ron is plausibly a victim of a relevant kind of inaccessibility, caused by his location in a HNIEC. This reading of Fricker's distinction is made further plausible when we understand Fricker's regard for USA1950s as implicitly treating it as a HNIEC. Recall Begby's definition of a HNIEC: where people "grow up in deeply prejudicial social settings, with no rational access to contrary evidence".[60] This seems to resemble Fricker's regard for Herbert's context in USA1950s. Furthermore, Fricker herself characterises Herbert in USA1950s as a subject of "bad luck", matching Begby's characterisation of those subject to HNIECS as victims

---

59. Full discussion of the metaphysical commitments of my argument is beyond the scope of this paper. For discussion of the relevant notion of "metaphysical trans-substantiation" (whereby "[a]lthough there is no new stuff in the world... new entities [can] come into being"), see Judith Baker, "The Metaphysical Construction of Value," Journal of Philosophy 86, no. 10 (1989): pp. 505-513, https://doi.org/10.5840/jphil1989861020, 507. For interesting discussions of the contingency of ethical judgements on metaphysical commitments, see Roy Sorensen, "Future Law: Prepunishment and the Causal Theory of Verdicts," Nous 40, no. 1 (2006): pp. 166-183, https://doi.org/10.1111/j.0029-4624.2006.00605.x and Roy Sorensen, "Future Law: Prepunishment and the Causal Theory of Verdicts," Nous 40, no. 1 (2006): pp. 166-183, https://doi.org/10.1111/j.0029-4624.2006.00605.x. My thanks to Dr. Alasdair Richmond for discussing such issues with me, and for these preliminary recommendations for further reading.

60. Begby, "Doxastic Morality", 168.

of bad luck.[61] In this way, USA1950s seems to be an example of a HNIEC.

On this basis, we have strong motivation for finding Fricker's analysis of USA1950s also applicable to RON, meaning that, by Fricker's own lights, RON is epistemically exculpated for the same reasons that Herbert in USA1950s is exculpated: Ron's failure to be testimonially virtuous with respect to the testimony of Indigenous Peoples is subject to the same exculpatory inaccessibility of conceptual resources as Herbert's. Thus we have a case of a contemporaneous prejudiced belief that meets the standards set by Fricker for epistemic exculpation. Were Ron to perpetrate a testimonial injustice against one of the Indigenous Peoples in the local reservation, for example, if they try to explain to him their history of being oppressed by white settlers, Ron, while at fault for perpetrating the epistemically unjust IPCD, would, on Fricker's own analysis, be epistemically exculpated. In such a case, as counterintuitive as it may seem, we would not be justified in having a "resentment of blame" towards Ron, but only a "resentment of disappointment" towards the circumstances that produced RON.

5   Conclusion

In closing, this paper has provided an epistemic basis for Cramer's imperative to view the prejudiced perspective of Ron, the logger from Northern Wisconsin, not as "just plain racism" but rather as the consequence of a much more complicated set of beliefs. In doing so, it has shown (i) that within the "politics of resentment", Ron's prejudiced perspective is more epistemically insidious than a common view of prejudice suggests, (ii) that Ron's location in a highly non-ideal epistemic context is sufficient for the epistemic exculpation of his prejudiced perspective, and (iii) that, in addition to historical-cultural distance, Ron's contemporary-cultural distance from the present moral discourse is sufficient for his epistemic exculpation. Taken together, these points suggest a counterintuitive, perhaps troubling conclusion about cases like RON. Insofar as an individual with an epistemically insidious belief has always been subject to a highly non-ideal epistemic context, any resentment of blame we harbour towards them is unjustified – we can only justifiably harbour a resentment of disappointment towards the circumstances that produced their prejudiced perspective.

References

Arpaly, Nomy. *Unprincipled Virtue: An Inquiry into Moral Agency. Oxford: Oxford University Press, 2011.*

---

61. Fricker, Epistemic Injustice, 101.

Arpaly, Nomy, and Timothy Schroeder. "Praise, Blame and the Whole Self." *Philosophical Studies 93, no. 2 (1999): 161–88. https://doi.org/10.1023/a:1004222928272.*

Baker, Judith. "The Metaphysical Construction of Value." *Journal of Philosophy 86, no. 10 (1989): 505–13. https://doi.org/10.5840/jphil1989861020.*

Begby, Endre. "Doxastic Morality." *Philosophical Topics 46, no. 1 (2018): 155–72. https://doi.org/10.5840/philtopics20184619.*

Begby, Endre. "The Epistemology of Prejudice." *Thought: A Journal of Philosophy 2, no. 2 (2013): 90–99. https://doi.org/10.1002/tht3.71.*

Bird, Michael Yellow. "What We Want to Be Called: Indigenous Peoples' Perspectives on Racial and Ethnic Identity Labels." *American Indian Quarterly 23, no. 2 (1999): 1. https://doi.org/10.2307/1185964.*

Bortolotti, Lisa, and Kengo Miyazono. "The Ethics of Delusional Belief." *Erkenntnis 81, no. 2 (2015): 275–96. https://doi.org/10.1007/s10670-015-9739-9.*

Cramer, Katherine J. *The Politics of Resentment: Rural Consciousness in Wisconsin and the Rise of Scott Walker. Chicago: University of Chicago Press, 2016.*

"Current Tribal Lands Map and Native Nations Facts." Wisconsin First Nations, March 31, 2020. https://wisconsinfirstnations.org/current-tribal-lands-map-native-nations-facts/.

Fricker, Miranda. *Epistemic Injustice: Power and the Ethics of Knowing. Oxford; New York: Oxford University Press, 2007.*

Fricker, Miranda. "The Relativism of Blame and Williams's Relativism of Distance." *Aristotelian Society Supplementary Volume 84, no. 1 (2010): 151–77. https://doi.org/10.1111/j.1467-8349.2010.00190.x.*

Goldberg, Sanford C. "Should Have Known." *Synthese 194, no. 8 (2015): 2863–94. https://doi.org/10.1007/s11229-015-0662-z.*

Keene, Adrienne. "Representations Matter: Serving Native Students in Higher Education." *Jcscore 1, no. 1 (2018): 101–11. https://doi.org/10.15763/issn.2642-2387.2015.1.1.101-111.*

Mason, Elinor. "Moral Ignorance and Blameworthiness." *Philosophical Studies 172, no. 11 (2015): 3037–57. https://doi.org/10.1007/s11098-015-0456-7.*

Robson, Jon. "A-Time to Die: A Growing Block Account of the Evil of Death." *Philosophia 42, no. 4 (2014): 911–25. https://doi.org/10.1007/s11406-014-9547-5.*

Sorensen, Roy. "Future Law: Prepunishment and the Causal Theory of Verdicts." *Nous 40, no. 1 (2006): 166–83. https://doi.org/10.1111/j.0029-4624.2006.00605.x.*

Does Metalinguistic Negotiation help with Conceptual Ethics?

Sophia Rommel*
University of St Andrews

Abstract Suppose a speaker states, "a fetus is a person," to which her conversational partner replies, "a fetus is not a person," and that their claims go beyond merely disagreeing (descriptively) about what the concept PERSON means. That is, the speakers are actually disagreeing (normatively) about what PERSON should mean, highlighting the dispute's normative, metalinguistic dimension. David Plunkett refers to this dispute as a 'metalinguistic negotiation'. There seems to be a gap between metalinguistic negotiation and the field of Conceptual Ethics, regarding the extent to which metalinguistic negotiation is connected to Conceptual Ethics. This missing link motivated me to explore if metalinguistic negotiation can help with Conceptual Ethics. I argue that the extent to which metalinguistic negotiation can help Conceptual Ethics, in terms of unifying Conceptual Ethics' projects to achieve shared goals, depends on which model of metalinguistic negotiation we utilise.

To explore my argument, I do six main things. First, I define Conceptual Ethics and offer an example of a defect concept. Second, I outline and discuss Plunkett's view, his model of metalinguistic negotiation, its strengths, and how it helps Conceptual Ethics. Third, I explore Plunkett's model's main weaknesses, identifying a dilemma. Fourth, I briefly consider and object to a suggestion attempting to salvage Plunkett's model. Fifth, I propose a reinterpreted model of metalinguistic negotiation, the Elasticity Model, utilising Speech Act Pluralism. In the proposal section, I introduce Speech Act Pluralism, bridge Speech Act Pluralism and metalinguistic negotiation, and outline the Elasticity Model. Thereupon, I investigate the Elasticity

---

*Sophia Rommel recently graduated with an MA (Hons) International Relations and Philosophy degree from the University of St Andrews. Sophia's main research interest in both International Relations and Philosophy is the complex relationship between language and practice. In Philosophy, this interest has propelled Sophia to focus on topics such as metalinguistic negotiation, speech acts, and conceptual ethics. In International Relations, the aforementioned interest shapes Sophia's approach to investigating United Nations peacekeeping missions in Sub-Saharan Africa, civilian protection, and globalisation. In the autumn of 2020, Sophia will start pursuing an MPhil in Politics and International Studies at the University of Cambridge as a Cambridge Trust Scholar.

Model's strengths and the ways in which it could help Conceptual Ethics. In particular, I attempt to demonstrate how the Elasticity Model can help Conceptual Ethics more than Plunkett's model can, particularly as the Elasticity Model avoids Plunkett's dilemma. Lastly, I raise and respond to numerous objections against my position.

1 Introduction

Suppose a speaker, S1[1], states, "Fs are G," to which her conversational partner, S2, replies, "Fs are not G," and that their claims go beyond merely disagreeing (descriptively) about what the concept G means. That is, the interlocutors are actually disagreeing (normatively) about what G should mean, highlighting the dispute's normative, metalinguistic dimension. To David Plunkett, this dispute is a metalinguistic negotiation (MLN)[2]. There seems to be a gap between MLN and the field of Conceptual Ethics[3], regarding the extent to which MLN is connected to Conceptual Ethics. This missing link motivated me to explore if MLN can help with Conceptual Ethics. Taking Conceptual Ethics to be a "sparse and scattered field,"[4] I believe it requires more unity to strengthen its "potential [. . . and help] conceptual development"[5]. I argue that the extent to which MLN[6] can help Conceptual Ethics, in terms of unifying Conceptual Ethics' projects to achieve shared goals, depends on which MLN model we utilise.

To explore my argument, I do six main things. First, I define Conceptual Ethics and offer an example of a defect concept. Second, I outline and discuss Plunkett's view, his MLN model, its strengths, and how it helps Conceptual Ethics. Then, I explore his model's main weaknesses, identifying a dilemma. Fourth, I briefly consider and object to a suggestion attempting to salvage Plunkett's model. Thereupon, I propose a different model of MLN, the Elasticity Model, utilising Speech Act Pluralism. Lastly, I raise and respond to several objections.

---

1. I refer to speakers as S1, S2, and so forth.

2. I explore Plunkett's model, as it is currently the most influential one.

3. I treat Conceptual Ethics and Conceptual Engineering as synonymous, as their definitions remain open questions. See Alexis Burgess et al., *Conceptual Engineering and Conceptual Ethics* (Oxford: Oxford University Press, 2020), 2 for further discussion.

4. Alexis Burgess and David Plunkett, "Conceptual Ethics I," *Philosophy Compass* 8, no. 12 (2013), 1096.

5. *Ibid.*, 1096-7.

6. I presuppose that MLNs exist and occur often.

2   Conceptual Ethics and Concepts

Conceptual Ethics asks, "which concepts should we use?"[7] In response, philosophers of Conceptual Ethics pursue diverse projects, which aim to improve our representational devices (such as concepts[8] and words, which arguably possess defects). Given the diversity of Conceptual Ethics, I define it broadly as "(i) [t]he assessment of representational devices [not only concepts], (ii) reflections on and proposal for how to improve representational devices, and (iii) efforts to implement the proposed improvements"[9]. In particular, concepts can have multidimensional defects, including semantic defects, socio-political defects, and inconsistency (enabling logical paradoxes), and one concept can entail numerous defects. Attempting to fix concepts is thus complex.

For example, consider MARRIAGE. Assuming[10] concepts have intensions (sets of principles picking out extensions) and extensions (references in the external world), the meaning of 'marriage' as commonly understood is often restricted to occurring only between heterosexual couples (indicating that MARRIAGE's intension includes only heterosexual couples). Therefore, MARRIAGE's intension picks out heterosexual couples, whilst excluding homosexual couples and polygamists. So, if used in policy, 'marriage' creates legal constraints on whom and how many people one can wed. For example, polygamy is illegal in the United States. Such socio-political defects require fixing, which is where a conceptual ethicist's task begins. She can attempt to engineer MARRIAGE by revising its meaning (revisionary conceptual engineering), replacing it with surrogates (replacement conceptual engineering), or eliminating it from our language (eliminativist conceptual engineering).

3   Plunkett's Approach to Metalinguistic Negotiation

Within Conceptual Ethics' realm, philosophers often seem to dispute what a concept means, creating a sense of disunity in the field and, if they are attempting to engineer the same concept, lacking cohesion between their projects. On Plunkett's view, such first-order, descriptive disputes concerning an expression's literal content (canonical disputes) are oftentimes not about literally expressed content (non-canonical disputes) and are metalinguistic. Plunkett argues that speakers can use "a word to communicate [their] views about the meaning of that word. These views can either be [about

---

7. David Plunkett, "Which Concepts Should We Use?: Metalinguistic Negotiations and The Methodology of Philosophy," *Inquiry* 58, no. 7-8 (2015).

8. I focus on concept-based Conceptual Ethics, as Plunkett does.

9. Herman Cappelen and David Plunkett, "A Guided Tour of Conceptual Engineering and Conceptual Ethics," in *Conceptual Engineering and Conceptual Ethics*, Alexis Burgess, Herman Cappelen, and David Plunkett (Oxford: Oxford University Press, 2020), 3.

10. I recognise the ongoing wider debates about concepts.

descriptive issues regarding] what a word does mean or about the normative issue of what it should mean"[11]. I return to Plunkett's focus on usage, first explaining the general descriptive/normative divide within the domain of non-canonical disputes. Plunkett distinguishes between metalinguistic disputes and MLNs, claiming that not all metalinguistic disputes are MLNs as some are purely descriptive (for example, those involving context-sensitive terms such as 'tall'[12]). Therefore, MLNs are a sub-set of metalinguistic disputes.

In order to therefore determine whether a dispute is a MLN, Plunkett proposes four main criteria. First, is it a dispute, namely, "a linguistic exchange that appears to express a disagreement"[13]? Second, does it truly express a disagreement? Third, do the speakers mean different things with (at least) one term? Fourth, is it normative? Plunkett qualifies his criteria with examples of evidence which can help pinpoint MLN. For example, Plunkett's third condition (C3) requires "good evidence that speakers in the dispute mean different things by (at least) one of the terms in that dispute"[14]. The 'good evidence' is that speakers consistently use the same term differently "in the same (non-defective) conditions"[15]. Regarding condition four (C4), Plunkett notes that the dispute is normative if it

> isn't just about descriptive information about what a word does mean, or how it is used. [. . .] [Evidence] here would be that speakers persist in their dispute even when they agree on the facts about a term's current meaning or current use.[16]

4   The General Strengths of Plunkett's Model

Initially, Plunkett's argument seems convincing. An immediate strength of Plunkett's view is his strong prima facie case for MLN's existence and frequent occurrence. In particular, Plunkett foregrounds disputes' possible normative dimension, which is an important feature of MLN. For example, S1's claim, "a fetus is a person," to which S2 counters, "you are mistaken, a fetus is not a person," may initially seem to concern the current meaning or use of 'person'. However, by applying Plunkett's criteria, one can deem the dispute normative as, arguably, the speakers are disputing what PERSON should mean, in virtue of observing their differing meanings of 'person' through their differing usages of 'person'. Due to normativity's defining importance for MLN, C4 is thus vital for MLN's existence.

---

11. Plunkett, "Which Concepts Should We Use?," 833.
12. *Ibid.*, 834-8.
13. *Ibid.*, 847.
14. *Ibid.*
15. *Ibid.*
16. *Ibid.*

Expanding upon my above claim, C4's emphasis on normativity highlights MLN's relation to Conceptual Ethics, entailing the potential for conceptual ethicists to explicitly engage in MLNs, thereby furthering their goals (for example, replacing MARRIAGE to improve social justice). C4 thus seemingly answers the title question, simultaneously offering to explain how: philosophers often engage in MLNs and, if aware of this (perhaps after looking for the specified evidence), they can realise what they are truly disputing and find a solution. As Plunkett and Sundell note,

> many disagreements about conceptual ethics are not expressed explicitly. Metalinguistic negotiations are the most important instances of this implicit [. . .] kind. [They] may not at first glance appear — either to the speakers themselves or to the theorist — to reflect disagreements about concept choice.[17]

Therefore, if disputing philosophers utilise Plunkett's criteria to reflect upon their disputes, they may find that their disputes express normative disagreements about concepts. So, is the missing link problem solved?

5   The Weaknesses of Plunkett's Model

Despite C4's strength, Plunkett's model faces two main problems: scope and an incompatibility. Regarding scope, all intractable disputes seem to be MLNs. Suppose S1 and S2 dispute whether or not a hedge is a tree, and that their dispute expresses a disagreement. Assuming their dispute persists beyond the facts about the current meaning or usage of 'hedge', their dispute could be classified as a MLN. Evidence for this claim is that, in this scenario, S1 uses 'hedge' to refer exclusively to shrubs, whilst S2 uses 'hedge' to refer exclusively to small trees. So, if such everyday disputes can be MLNs on Plunkett's view, how many other disputes are MLNs? Plunkett simply states that discerning this is "no easy task"[18].

I return to the scope worry later, focusing my analysis on an incompatibility between C3 and C4, which I argue undermines Plunkett's model's coherence and, subsequently, its ability to help Conceptual Ethics. Whilst C3 specifies that speakers can mean different things with a term, evidenced by their differing usages of it[19], C4's evidence (for normativity) is that "speakers persist in their dispute even when they agree

---

17. David Plunkett and Tim Sundell, "Disagreement and the Semantics of Normative and Evaluative Terms," *Philosophers' Imprint* 13, no. 23 (2013), 3.

18. Plunkett, "Which Concepts Should We Use?," 850.

19. Plunkett seemingly adheres to semantic holism (every change in use changes meaning), characterising language as hyper-sensitive to usage change. Thanks to Dr Patrick Greenough for highlighting this point in a discussion.

on the facts about a term's current meaning or current use"[20]. To Plunkett, the facts are either about their usage(s) or their speech community's, and the speakers negotiate "how they should use the term going forward,"[21] simpliciter. However, I wonder how the speakers can agree on the facts about a given term's current meaning or use (hereupon, current meaning and use facts are referred to as 'the facts'), given that they can mean different things with the term. I believe this incompatibility arises from Plunkett's overly strong Use Theory of Meaning. Whereas the Use Theory of Meaning generally claims that use determines meaning, Plunkett's aforementioned conditions imply a bi-entailment: use determines meaning and any difference in meaning is a difference in use. This bi-entailment seems problematic for two main reasons.

First, as C4 allows speakers to agree on their usages' divergence, Plunkett implicitly suggests use facts[22] are intrinsically normative rather than descriptive, as deciding upon a term's meaning would involve selecting one speaker's use over another's. Given bi-entailment, this also suggests meaning facts are intrinsically normative. For example, utilising Plunkett's linguist example, a linguist can state F "just means"[23] G, given what G is (descriptively). This implies that in a MLN, an expert on terms, such as a linguist, will have their usage of a term accepted and a non-expert's usage will be rejected, so the facts ultimately depend on who states them. By extension, considering Plunkett's model's implications for Conceptual Ethics, as C3 permits meaning difference and Plunkett neither specifies exit rules nor how to arrive at a solution (deciding which concept to use), his model seems impractical if we consider unifying projects more towards common goals, as disputes and disunity can indefinitely continue.

Second, considering a dispute about PERSON, C3 and C4 allow S1 and S2 to agree that S1 uses 'person' to apply to fetuses and S2 does not. If their dispute persists beyond such agreement, it is a MLN, supporting Plunkett's position. Given C4's stipulation, Plunkett would conclude from this that the speakers are negotiating what PERSON should mean simpliciter, rather than what PERSON should mean to each individual. However, in order to claim the former, Plunkett would need to allow for 'person' to possess a shared social (public) meaning (which is then disputed in the MLN), yet C3 does not allow for this and C4 stipulates that the facts (such as speakers' diverging usages) about 'person' must be agreed upon to make their dispute normative.

To clarify my aforementioned point, as the speakers can agree they mean different things with 'person' (individually), how can Plunkett claim they are disputing what 'person' means (simpliciter)? Plunkett overlooks the possibility that if speakers agree that their usages diverge, their agreement does not equate to one on a term's (general) current meaning and use facts. My objection creates a dilemma for Plunkett: if he adheres to C3, he loses C4 (and cannot bridge the gap to Conceptual Ethics) as there

---

20. Plunkett, "Which Concepts Should We Use?," 847.
21. *Ibid.*, 840.
22. I wonder if this applies to all use facts or only to some.
23. Plunkett, "Which Concepts Should We Use?," 842.

arguably cannot be agreement on the facts. If he keeps C4, he cannot advocate meaning difference (key to his model). Thus, although C4 is crucial for MLN, Plunkett cannot address the missing link fully. Therefore, we can either fix Plunkett's model or propose a meaning sameness model.

6   Salvaging Plunkett's Model?

If we attempt to salvage Plunkett's view, one possible approach is Simple Contextualism as, on one reading, meaning difference could imply terms' context-sensitivity. MLNs concern "how to best use a word relative to a context,"[24] which supports this connection. Simple Contextualism claims that speakers share a term's conventional linguistic meaning. A conventional linguistic meaning is incomplete[25], so a given context's features complete a term's semantic value. Therefore, one could argue that requiring context to fix content accounts for a context-sensitive meaning difference notion, implying a reliance on contexts to derive complete meanings. Therefore, a conventional linguistic meaning without a specified context can create the impression of speakers meaning different things with a term, accommodating for C3.

Whilst an intriguing suggestion, I object to it on two fronts: first, if someone made the above argument, they would overlook Simple Contextualism's single-content view. Simple Contextualism is therefore actually an attractive option for meaning sameness rather than meaning difference, suggesting a possible single-content contextualist MLN. That is, contrasting Plunkett, propounding meaning sameness arguably accounts for C4 (and, thus, normativity), bridging the missing link and solving Plunkett's dilemma as speakers can agree or disagree about what a word means[26], since one meaning exists. Notwithstanding, this approach may produce a dilemma for Plunkett insofar as that he could not propound meaning difference. In conjunction, as Plunkett claims context-sensitive terms are insufficiently normative (for example, 'tall'), he may reject Simple Contextualism as a viable alternative. Second, even if I utilised Simple Contextualism to construct a meaning sameness model, I would be committed to claiming that all words are context-sensitive. Considering the evidence suggesting the contrary, I find this too ambitious a claim to defend here, particularly as I seek to find a model entailing minimal baggage.

Plunkett's dilemma and Simple Contextualism's limitations lead me to explore my second option: proposing a meaning sameness model, the Elasticity Model, aiming to avoid Plunkett's problems. An important preliminary is that I do not intend the Elasti-

---

24. Plunkett and Sundell, "Disagreement and the Semantics of Normative and Evaluative Terms," 3.

25. Joseph Almog et al., *Themes from Kaplan* (Oxford: Oxford University Press, 1989) discusses Kaplan's character and content distinction.

26. Plunkett could argue that this captures canonical disputes. However, I think that a 'simple contextualist MLN' could demonstrate how terms' incomplete meanings elongate disputes.

city Model to be a flawless model. Adhering to my paper's purpose, I simply explore how the Elasticity Model could help Conceptual Ethics more than Plunkett's model. Therefore, I consider some objections after presenting the Elasticity Model, which indicates that the Elasticity Model requires more conceptual development than I can offer in the space of my paper. To briefly outline the Elasticity Model's main claim, it holds that interlocutors in a MLN are not directly negotiating about what our concepts should mean, rather, they are negotiating what their uttered sentences (embedding concepts) about concepts should be saying. More specifically, speakers are negotiating which expressed proposition they should associate with an uttered sentence[27]. I elaborate on the Elasticity Model after introducing its theoretical background (Speech Act Pluralism).

7   Speech Act Pluralism

I utilise Speech Act Pluralism for MLN as I find Plunkett's focus on usage important because speakers use the terms they utter in their disputes. Speech Act Pluralism helps me foreground and maintain a focus on usage, albeit differently to Plunkett, as Speech Act Pluralism claims "no one thing is said [. . .] by an utterance; rather, indefinitely many propositions[28] are said, asserted[29], claimed, stated"[30]. Therefore, what is said when a speaker utters a sentence (for example, of the form 'X is F') in a context of utterance expresses multiple propositions (P1, P2, P3, and so forth). Speech Act Pluralism thus distinguishes between a sentence's literal meaning and speaker meaning (namely, "what a person means in using [a sentence]"[31]).

For example, Speech Act Pluralism[32] would claim that a speaker asserting "Trump is the US President" in a context of utterance expresses multiple propositions, such as *Donald is the US President*, *Melania's husband is our President*, *The Trump Organisation's owner is the President*, and so forth, due to a one-many saying relation. Specifically, the saying relation, relativised to a single context, is one-many. Moreover, "pluralism about speech act content applies [both] to sentences that report something about what speakers' utterances say [and] to the reported speakers' utterances,"[33] which I expand upon below.

---

27. For my paper's purposes, I only refer to the utterances of sentences.

28. Every expressed proposition has truth conditions (necessary and sufficient conditions), illustrated below.

29. Speakers produce speech acts by, for example, asserting, "Fs are G."

30. Herman Cappelen and Ernest Lepore, *Insensitive Semantics: a Defense of Semantic Minimalism and Speech Act Pluralism* (Malden, MA: Blackwell, 2007), 4.

31. Gillian Russell and Delia Graff Fara, *The Routledge Companion to Philosophy of Language* (New York: Routledge, 2015), 79.

32. Although this view's accuracy is contested, I believe that the way we utter sentences and use reports in natural language supports Speech Act Pluralism. I aim to demonstrate this through my examples.

33. Cappelen and Lepore, *Insensitive Semantics,* 199.

Aside from an interlocutor's utterance in a context of utterance, also referred to as the reported context (what is said), Speech Act Pluralism considers "the context of those who say or think about what the speaker said, [so] the context of those who report what's said by the utterance can [partly] determine what was said by that utterance"[34]. Therefore, if S1 asserts, "X is F," this "does not stand in a single one:one relationship with some indirect speech report [(hereupon, 'report')], but rather explodes into a plethora of possible indirect speech acts"[35]. So, if S2 reports on S1's utterance, S2's report selects one of its expressed propositions. Therefore, the expressed propositions (multiple sayings), which expressed proposition a speaker reports on, and what is saliently said depend partly "on features quite external to that [uttered] sentence, like the [given] conversational context [. . .] and the (shared) background assumptions of speakers and [hearers]"[36]. Some reports may therefore communicate (in the reporting context) "things only tangentially connected with the meaning of the sentence produced,"[37] and which expressed proposition is most felicitous to report on ultimately depends on the speakers' freedom to select one. Thus, both utterances' and reports' flexibility stem from the one-many saying relation's flexibility[38].

8   Initial Worry

Initially, one may perceive a fundamental problem with Speech Act Pluralism, namely, why are some reports successful and others not? Is this context-sensitivity? Does Speech Act Pluralism share Simple Contextualism's limitations? I contend that one can distinguish between the two views. Whilst Speech Act Pluralism proposes that an utterance produced in a context of utterance expresses multiple propositions, Simple Contextualism focuses on terms' context-sensitivity and how contexts complete terms' meanings. Figure 1, which I have produced, underscores this difference:

---

34. *Ibid.*, 4-5.

35. Emma Borg, *Minimal Semantics* (Oxford: Oxford University Press, 2004), 114-5.

36. *Ibid.*, 115.

37. *Ibid.*

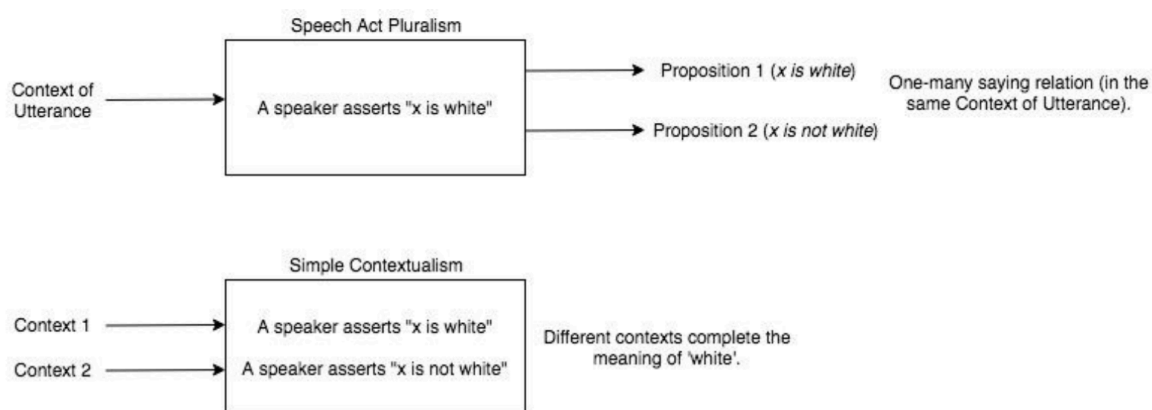38. The 'elasticity' of 'Elasticity Model' denotes this flexibility.

Figure 1: Speech Act Pluralism's and Simple Contextualism's differences regarding context

9   Using Speech Act Pluralism for Metalinguistic Negotiation: the Elasticity Model

With the Elasticity Model's background established and the contextualism worry dispelled, I now explore how MLN could work on a Speech Act Pluralism view. First, I propose its MLN criteria: are the interlocutors engaged in a dispute? Does it express a disagreement? Is there evidence that the interlocutors mean the same thing with a term used in a sentence and are disputing the uttered sentence's expressed propositions (EC3)? Is it normative (EC4)? My criteria suggest MLNs may not always overtly involve disputes (in Plunkett's sense) as the expressed propositions, rather than a term's meaning, are multiple, and speakers must select what is saliently said among the many sayings[39]. Considering Plunkett's Secretariat example, wherein S1 says, "Secretariat is an athlete," and S2 disagrees, Plunkett argues that the speakers are negotiating what ATHLETE should mean. For S1, non-human animals can be athletes, whereas, for S2, they cannot, and this meaning difference is evidenced by S1 using 'athlete' to refer to Secretariat, a horse, whereas S2 does not. In contrast, the Elasticity Model views the speakers as negotiating what "Secretariat is an athlete" should be saying (what the most relevant expressed proposition is for their MLN). Thus, 'athlete' possesses a single (shared) meaning and S1's utterance of "Secretariat is an athlete," with respect to a context of utterance, expresses multiple, differing propositions. Moreover, an utterance's multiple expressed propositions can account for how interlocutors seem to mean different things when they utter a sentence using the same (disputed) term, accounting for an initial impression of meaning difference.

---

39. What is said does not imply that speakers simply talk past each other as, despite shared semantic-content, disputes persist.

Nonetheless, one may question whether the interlocutors are aware a MLN is occurring. Moreover, how can speakers decide upon what is said given this complicated view and that there is "no one correct answer to what was said by an utterance, [nor. . .] to what was said by a report of [the utterance]"[40]? I suggest that a filtering mechanism is needed, such as MLN referees. Arguably, if interlocutors are aware of engaging in MLNs, they could referee themselves, provided they know an uttered sentence has multiple expressed propositions. Referees could choose the most felicitous saying as "intuitively we don't want to countenance all of them as semantically informative,"[41] giving the referees a demanding decision-making role. I address some practical questions about such decision-making below.

10   The Elasticity Model and Conceptual Ethics

To test the Elasticity Model, I consider an example[42] pertaining to Conceptual Ethics and connect it to Conceptual Ethics projects. Suppose S1 utters, "x is white" in a context of utterance, expressing propositions P1, P2, and so on. We could think of the truth conditions for 'x is white' as ''x is white' is true if and only if (iff) x has a reflectance property greater than or equal to ($\geq$) n'. So, P1's truth condition is 'x is white iff x has a reflectance property $\geq$ 0', whereas P2's is 'x is white iff x has a reflectance property $\geq$ 1', and different speakers can report on different expressed propositions, "all of which may be licensed by [S1's] production of ["x is white"]"[43]. The speakers are thus negotiating the boundary for whiteness and, considering EC3, there is shared meaning insofar as what S1 and S2 mean by WHITE is the same, however, their choice of what is saliently said differs.

For example, S1 argues, "x is white," satisfying P1's truth conditions (a lower standard), whereas S2 does not think x is white because it must satisfy P2's truth conditions (a higher standard). Regarding EC4, one may argue that MLNs express merely verbal disputes as 'white' has one meaning, so the speakers could be talking past each other. However, I argue that normativity is present, as what should be saliently said is normatively disputed. That is, the Elasticity Model proposes a MLN that concerns which of the many truth conditions we should associate with a sentence, in a context of utterance. Thus, in certain contexts, S1 and S2 can agree because of their shared meaning of WHITE and by deciding what an utterance about WHITE should say for their MLN's purposes. It is a further, practical question as to what the most felicitous saying is, which suggests setting standards and artificial constraints on these, depending on the MLN. Arguably, speakers must agree upon which standards are appropriate in which condition(s), which I explore below.

---

40. Cappelen and Lepore, *Insensitive Semantics*, 199.
41. Borg, *Minimal Semantics*, 115.
42. As I aim to challenge Plunkett's account, I too do not consider concepts such as TRUTH.
43. Borg, *Minimal Semantics*, 114-5.

Applying my example to a wider discussion wherein theorists consider normative issues about our conceptual choices of racial terms, "x is white" may hold if the theorists adopt a low standard (perhaps for everyday discussions involving WHITE). However, if engaging in an eliminativist project on WHITE, demanding standards may be adopted for the MLN, whereby "x is white" does not hold. For example, Appiah, an eliminativist about RACE, advocates that "there are no races"[44] as no referents exist for it. Appiah claims "the evil [...] is done by the concept,"[45] deeming RACE defective, especially when racial terms are used in policies to oppress certain groups of people. An Appiah-like eliminativist would thus argue that racial terms should not be used, as they do not refer to anything. In contrast, racial constructivists perceive races as socially constructed, leading many to argue that we should keep RACE (and racial terms), as removing it leaves out "something [...] causally or socially important"[46]. For example, Root argues that labelling races through terms such as 'white' and 'black' is "conceptually or logically required"[47] to understand the (oftentimes harmful) role these terms have played and continue to play in societies. Therefore, theorists such as Root hold that 'race' and "racial terms pick out groups of persons in virtue of either superficial or culturally local features (or both),"[48] contrasting Appiah's claims.

Utilising the aforementioned theorists' diverging perspectives, suppose S1 from my WHITE example is a Root-like theorist and S2 is a Appiah-like theorist, and both desire to further social justice in different ways. Whereas S1 aims to revise WHITE's intension, S2 holds that WHITE is so defective that it must be removed. Given Speech Act Pluralism, if S1 argues, "Jim is white," her utterance expresses multiple propositions (depending on what is present in the context of utterance and reporting context) including *Jim is not white*[49], which S2 may decide is saliently said (given S2's views). Thus, which expressed proposition either speaker reports about the other's utterance whilst refereeing depends on what they deem most felicitous to report. Furthermore, the interlocutors could decide what is saliently said in their MLN by setting standards. That is, they decide which expressed proposition (and its truth conditions) are correct for an utterance about WHITE for their MLN's purposes[50]. Consequently, their dispute may be long and confusing[51]. By extension, regarding my example, the Elasticity Model highlights that if a revisionist and an eliminativist agree on a standard, this may impact one of their projects more than the other's. For example, if they agree "Jim is

---

44. Kwame Anthony Appiah, "The Uncompleted Argument: Du Bois and the Illusion of Race," *Critical Inquiry* 12, no. 1 (1985), 35.

45. *Ibid.*

46. Ron Mallon, "'Race': Normative, Not Metaphysical or Semantic," *Ethics* 116, no. 3 (2006), 534.

47. *Ibid.*, 536.

48. *Ibid.*

49. Some Speech Act Pluralism theorists find arguing that one uttered sentence, as used in a single context, can express incompatible propositions controversial, whilst others remain neutral on this matter.

50. Essentially, they decide how to categorise our world correctly, implying a deflated sense of normativity.

51. This is quite representative of our drawn-out disputes.

white" expresses *Jim is white*, the eliminativist may need to rethink his project[52].

Arguably, my above example demonstrates the Elasticity Model's applicability, as the interlocutors mean same thing with WHITE and are disputing what the utterance "Jim is white" should be saying, rather than what WHITE's meaning should be (fulfilling EC3). Considering EC4, I view this dispute as normative because the interlocutors are disputing what the uttered sentence should be saying, not what it means or says. Therefore, I believe the example highlights that the Elasticity Model can be applied to Conceptual Ethics-style disputes, ridding MLN of Plunkett's dilemma (as there is a shared meaning of 'white') and Simple Contextualism's limitations (meanings are not incomplete). Moreover, the Elasticity Model can bridge MLN and Conceptual Ethics as, contrasting Plunkett's lack of exit rules, standard-setting enables a greater chance for speakers to arrive at a solution. That is, as Conceptual Ethics asks, "which concepts should we use?" the Elasticity Model indicates multiple ways to decide upon which concepts to use, depending on the operative standards in a MLN's context. Therefore, the Elasticity Model arguably helps Conceptual Ethics more than Plunkett's model, as it possesses three advantages: coherence, unity, and practicality. As I demonstrated, the Elasticity Model can unify projects towards common goals, such as furthering social justice, because interlocutors can find solutions. Thus, for the "sparse and scattered field"[53] of Conceptual Ethics, the unity the Elasticity Model could provide, especially if further developed, is a virtue.

11   Objections and Replies

The Elasticity Model faces numerous challenges, which the below objections highlight. Moreover, these objections are not exhaustive, so I welcome additional comments for refinement.

What if different MLN groups decide upon a different expressed proposition for an utterance?

The groups need to engage in MLN with each other, in order to re-negotiate an utterance's expressed proposition (and, thus, what is saliently said in their new MLN). Hopefully, their MLN can enable them to eventually converge on one expressed proposition.

What if an interlocutor sticks to another expressed proposition?

Perhaps she should not participate in MLNs. As Plunkett notes, being aware of engaging in MLNs is rare, so a willingness to explicitly participate in, and accept, MLN

---

52. Making such choices may reduce projects' diversity, a potentially negative implication.
53. Burgess and Plunkett, "Conceptual Ethics I," 1096.

is arguably required. The reluctance to agree with a solution underscores the need for negotiation rules, especially if MLN becomes more popular (as a research interest) and accepted as a phenomenon (philosophically and socially).

> *The Elasticity Model faces Plunkett's scope issues.*

I agree and bite the bullet: perhaps most disputes are MLNs. I agree with Plunkett that determining MLN's scope is difficult, especially if we are still developing more coherent criteria. Nonetheless, I hope to have shown how the Elasticity Model avoids Plunkett's dilemma, which I identified as the major weakness of Plunkett's model.

> *Does the Elasticity Model's speech act pluralist-nature still align it with Conceptual Ethics' question of "which concepts should we use?"*

Yes. For the Elasticity Model, the above question is asked in virtue of an utterance's expressed propositions (embedding concepts) and which of these holds depends on a given MLN's purpose(s).

> *What about scenarios exploring an utterance's indefinite expressed propositions? Is this an endless dispute?*

The Elasticity Model needs more conceptual work to answer this question properly. However, if one adheres to a speech act pluralist approach, perhaps MLNs are just incredibly intractable (therefore, also wide in scope) because of the one-many saying relation. Thence, we may need to simplify and choose for the purposes of our negotiations.

> *Speech Act Pluralism just gives us what a sentence is saying, not what a sentence should be saying. So, the Elasticity Model is insufficiently normative.*

This objection initially makes a good point; however, it overlooks that what is saliently said is partly determined by normative matters concerning which expressed propositions we should be using. Thus, I believe the Elasticity Model is sufficiently normative.

12　Conclusion

I have argued that the extent to which MLN can help Conceptual Ethics, in terms of unifying Conceptual Ethics' projects to achieve shared goals, depends on which MLN model we utilise. I introduced Conceptual Ethics and discussed Plunkett's MLN criteria. Although I identified C4 as crucial for MLN, as MLN's normativity connects MLN to Conceptual Ethics and highlights our disputes' (potentially) normative dimen-

sion, I raised two objections and analysed their implications through examples, focusing on Plunkett's dilemma. Prior to discarding Plunkett's model, I considered tackling the dilemma through Simple Contextualism. However, I found Simple Contextualism more helpful for a meaning sameness rather than a meaning difference model. As using Simple Contextualism for MLN would force me to deem all words context-sensitive, I rejected this idea and proposed the Elasticity Model.

First, I qualified my proposal and introduced the Elasticity Model's Speech Act Pluralism background. Thereafter, I demonstrated how the Elasticity Model could help Conceptual Ethics more than Plunkett's model, including avoiding its limitations and solving the missing link. Through a key example, I found the Elasticity Model to entail the potential to help Conceptual Ethics projects, especially those on the same concept, unify more towards goals philosophers may share, albeit with numerous limitations. I then considered some of these limitations as objections. Overall, I hope to have presented the Elasticity Model as a viable starting-point for developing a more coherent MLN model.

References

Almog, Joseph, John Perry, and Howard Wettstein, eds. *Themes from Kaplan*. Oxford: Oxford University Press, 1989.

Appiah, Kwame Anthony. "The Uncompleted Argument: Du Bois and the Illusion of Race." *Critical Inquiry* 12, no. 1 (1985): 21-37. https://www.journals.uchicago.edu/doi/10.1086/448319.

Borg, Emma. *Minimal Semantics*. Oxford: Oxford University Press, 2004.

Burgess, Alexis, and David Plunkett. "Conceptual Ethics I." *Philosophy Compass* 8, no. 12 (2013): 1091–1101. https://doi.org/10.1111/phc3.12086.

Burgess, Alexis, Herman Cappelen, and David Plunkett. *Conceptual Engineering and Conceptual Ethics*. Oxford: Oxford University Press, 2020.

Cappelen, Herman, and David Plunkett. "A Guided Tour of Conceptual Engineering and Conceptual Ethics," in *Conceptual Engineering and Conceptual Ethics*, Burgess, Alexis, Herman Cappelen, and David Plunkett. Oxford: Oxford University Press, 2020.

Cappelen, Herman, and Ernest Lepore. *Insensitive Semantics: a Defense of Semantic Minimalism and Speech Act Pluralism*. Malden, MA: Blackwell, 2007.

Mallon, Ron. "'Race': Normative, Not Metaphysical or Semantic." *Ethics* 116, no. 3 (2006): 525-551. https://www.journals.uchicago.edu/doi/10.1086/500495.

Plunkett, David. "Which Concepts Should We Use?: Metalinguistic Negotiations and The Methodology of Philosophy." *Inquiry* 58, no. 7-8 (2015): 828-874. https://www.tandfonline.com/doi/full/10.1080/0020174X.2015.1080184.

Plunkett, David, and Tim Sundell. "Disagreement and the Semantics of Normative and Evaluative Terms." *Philosophers' Imprint* 13, no. 23 (2013): 1-37. https://quod.lib.umich.edu/p/phimp/3521354.0013.023/1.

Russell, Gillian, and Delia Graff Fara. *The Routledge Companion to Philosophy of Language*. New York: Routledge, 2015.