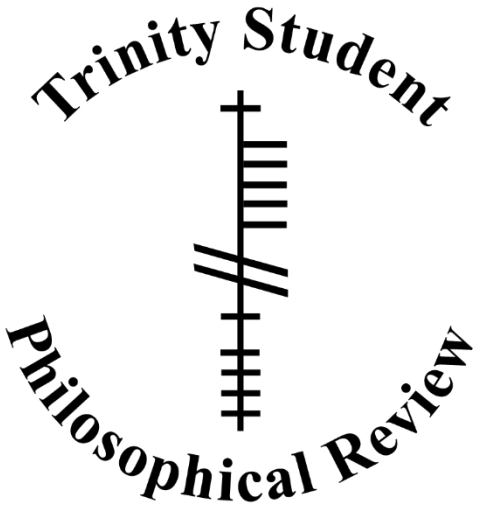
Trinity Student Philosophical Review

Vol. III

Trinity College Dublin 2025



Patron

Professor John Divers

Head of the Philosophy Department, Trinity College Dublin

Published in March 2025 by the Trinity Student Philosophical Review,
c/o Department of Philosophy,

Trinity College Dublin, Dublin 2, Ireland

For an online version please visit: https://www.tcd.ie/philosophy/current-students/trinity-student-philosophical-review/

A catalogue record for this journal is available from Trinity College Library, Dublin.

Printed by printbureau.ie

Cover art by Helena Cutbill

The Trinity Student Philosophical Review would like to thank all our sponsors:

Tethras Technology Ltd.

printbureau.ie

Trinity Student Philosophical Review Committee – Volume III

Editor-in-Chief
Fiona Lüling
luelingf@tcd.ie

Editor
Lewis Sharp
sharpl@tcd.ie

Editor
Ryszard Madejczyk
madejczr@tcd.ie

Events Manager
Giulia Villa
villag@tcd.ie

Social Media Manager
Ilya Semple
semplei@tcd.ie

Treasurer
Georgina Dowling
dowling1@tcd.ie

Winner of the Iris Murdoch Prize

for the best undergraduate Essay 2025

First-Person Perspective(s) and Subjectivity: A Different Problem of Other Minds

By Giorgia Carli

Editorial Prelude:

Pausing for thought — this is the goal of Volume III of the Trinity Student Philosophical Review (TSPR). I am delighted to welcome you to seven pieces that were selected to represent the scholarly excellence of Trinity's undergraduate students. This review is designed to offer a moment of reflection—a brief pause from the rush of everyday life—to engage with thought-provoking philosophical discussion. Founded just three years ago, the TSPR was established to support philosophy students in their academic endeavours and to showcase their original thoughts and critical argumentation. With each volume, we strive to present a diverse selection of essays, ensuring that every reader can find something to enjoy. Each essay is introduced by a short introductory paragraph, written by my incredible editorial team, to allow each reader, with or without a philosophical background, to gain a preliminary understanding of the essay that follows.

On opening this booklet, the reader may wonder what the symbol in the TSPR logo and on the pages of this issue means. It is Ogham, the earliest form of writing found in Ireland. Ogham, with its lines and notches, is typically found on memorial stones throughout Ireland. Characterised by being written vertically and read from bottom to top, our chosen symbol can be translated as 'wisdom'. Wisdom is demonstrated by going beyond mere knowledge. Wisdom reshapes, redefines and perfects a thought until it becomes an argument— an idea that resonates deeply with the purpose of this publication.

With each edition of the TSPR, we also seek to recognize excellence by awarding a prize to the best essay submitted. Until now, this prize bore the name of Edmund Burke. However, inspired by Trinity's renaming of the Boland Library and the need to confront the moral complexities of historical figures, we have chosen to rename it the ***Iris Murdoch Prize***.

Born in Phibsborough, Dublin, and later educated in England, Iris Murdoch studied at Oxford before pursuing philosophy under Ludwig Wittgenstein at Cambridge. Murdoch, besides being a philosopher, is best known as a novelist. Freedom and love are central themes in her novels. Murdoch believed that true freedom requires stripping away self-centred desires and allowing characters to act naturally, free from the impositions of the author. This created freedom leads to her books being characterized by ambiguity, dark humour, comic, but also realism, especially when it comes to her detailed description of her characters. Murdoch's philosophy is also represented in her fiction in the way that our world cannot be detangled from the fluid transition from "the fantastic and the ordinary, the plain and the symbolic". For her influence and impact throughout literature, but also for her representation of philosophy in a modern way, we have chosen to name the prize after her legacy.

For an essay that is brilliantly argued, rich in detail, and leaves the reader in a state of thoughtful reflection and puzzlement about our world, we are delighted to award the Iris Murdoch Prize of the *Trinity Student Philosophical Review* to **Giorgia Carli**!

The applicability of the phenomenon described in Carli's outstanding essay led us to choose her to start of this year's edition. When we think of our seemingly common understanding of things like hunger and thirst, shame, itchiness or fatigue, we all seem to have a clear objective understanding of them. Moreover, we may even have a definition in mind, but this definition, Carli argues, is significantly shaped by our individual experience of it and cannot be regarded as a general objective understanding at all. Carli explores this problem in a very clear way, using the example of pain to guide the reader through her argument.

Following the theme of privacy in our understanding of the world, **Nina Stremersch's** essay takes on the well-known showing/saying distinction by Ludwig Wittgenstein. Stremersch argues for the need for meaningfulness in our world, which can only be found in speech acts through a public and objective language.

This nuanced essay is followed by a piece by **Miriam Treitinger**, who takes us to the famously hard to digest Immanuel Kant. However, she presents his ideas in a lucid way and enlightens us about the importance of schemata in our thought. Treitinger presents schemata as an essential part of our thought and image-producing mental processes. One question she raises is how we can determine causality from our intuition that it was I who pulled the trigger and made the gun shoot?

Shooting the gun can be something immoral if the target is something of innate value. Emotivism would present itself as a theory that evaluates such morality of actions by the accompanying attitudes of approval or rejection. **Eoin Moran** rejects emotivism all together in a brilliant way, by combining not one objection but two in their shared force against it.

Moran's essay poses significant challenges to the emotivist theory, leading us to **Robert Finan**'s piece, which follows with a cohesive evaluation of a thought-provoking mathematical puzzle. Can infinity be differently big? Say you subtract one from one infinity and add one to another, both are surely still infinite yet the sets of numbers within infinity would therefore have different values. Finan leads us through the process of unpacking this problem in a profound way.

With a clearer understanding of infinity and cardinality, we must return to a problem that underlies the question we posed earlier. If the gun which was fired earlier had missed its target, would it have hit anything else? The questions about possibility are called counterfactuals. **Cathal Lacey** explores the relationship between counterfactuals and necessity in a compelling way.

Last but not least, we round this year's edition up with a "Gaumenschmauss" as we say in Germany, or a culinary delight in English, which does not quite capture the same essence. Regardless, **Gavin Dunphy** brings us back to reality and leaves us puzzled about what we know about the non-existent. How many ends does a hole in a cheese have? The crack in my

mother's favourite vase, is it a crack that exists or the non-existence of the vase? How can we talk about the existence of the crack if its property is not to be?

While the reader may be left in a puzzled state, after reading through seven thought-provoking essays, I hope that we have touched upon the many facets of philosophy and how it engages us in thought. As an editor-in-chief I immensely enjoyed curating a review about the ideas that keep us thinking long after we have read the last words in a piece of philosophical exquisite. It is philosophy that seeks for truths in the plurality of ideas. It is trying to make a clear thought out of the jumble in which we find ourselves.

This idea is captured in the cover artwork by the artist and my dear friend Helena Cutbill. Philosophers try to reach solutions that seem close but remain far away from what we know. Our thoughts need time to untangle and form a clearer picture. Philosophy navigates us to that enlightened stage through logical reasoning. This itself is an art represented by the essays in this volume.

On this remark I wish each and every reader patience, enjoyment and provocative thoughts reading the essays of this year's volume.

This incredible edition would not have been possible without the help of many people. Of course, the team behind the TSPR has been working restlessly the past weeks and months to make it possible for us to hold this copy in our hands today. Thank you for your trust and patience! Equally important has been the help of the Philosophy Department and especially our Patron John Divers. Particularly, I would like to thank Katy Armstrong, Carly Forde and Farbod Akhlaghi for always being so helpful. With patience and effort, all our questions were answered and solutions were found. Furthermore, I would like to thank our sponsor Tethras Technology and printbureau for their generosity. Finally, before you can dwell on the next 60 pages of philosophical discourse, I would like to thank all contributors. Without your contribution this would not have been possible. Now I am left to remind you that this edition is here for you to pause for some thought and to remember that:

"A serious and good philosophical work could be written consisting entirely of jokes." (Ludwig Wittgenstein)

Fiona Lüling

Editor-in-Chief TSPR

Table of Content

First-Person Perspective(s) and Subjectivity: A Different Problem of Other Minds

By Giorgia Carli

We all experience our lives in very different ways. Yet we are able to communicate about our experiences as if they were shared by everyone. It is through speech acts that we confirm and consolidate our shared understanding of the world. Does that mean it becomes objectively the same? Carli takes the example of pain and how pain can only be understood through one's own experience. Carli explores this topic in such a sophisticated way that it makes you rethink your perception and interaction with the world.

This paper aims at examining McGinn's *conceptual problem of other minds* (CP). In section 1, I will provide a broad definition of CP, and then proceed to analyse how, for McGinn specifically, it issues a difficulty in our formulation of a univocal concept of pain. In section 2, I will disclose an internal inconsistency in McGinn's appeal to behaviourism. I will then advance my own thesis, arguing firstly that we view others as being first-person, rather than third-person, perspectives, and that this consequently warrants a univocal use of our first-personal concept of pain. In Section 3, I will consider two objections against my thesis. The first objection formulates a new version of CP as a challenge in our understanding of first-person perspectives different from our own. I will reply to this new CP through Donald Davidson's three-fold form of empirical knowledge, understood as subjectivity, alter-subjectivity, and objectivity. The second challenge alleges a self-refuting dilemma in my answer to the new CP, which arises from an inconsistency between the terms "first-person perspective" and "subjectivity". I will refute this second objection by recalling their corresponding meanings.

1. McGinn's Problem of Other Minds

Broadly speaking, the *conceptual problem of other minds* (CP) states that there are obstacles in the attribution of a general concept of mind, which weaken its very generality (Avramides, 2001, p. 220). However, we seem to have a general concept of mind (Gomes, 2011, p. 356) – that is, a concept of mental life that can be ascribed to multiple distinguishable individuals (Avramides, 2001, p. 220). The question then raised regards the kind of obstacles that these are, and it is subject to extensive research in the literature.

On this topic, Colin McGinn offers a notable contribution in his article 'What Is The Problem of Other Minds?' (1984, pp. 119-137), in which he argues that CP issues a difficulty in our abstraction of a neutral conception of experiences that we inevitably undergo through a first-person perspective (*id.*, p. 136). McGinn starts his research by observing that there exists an uncommon element associated with one's own experiences, which he terms "introspection" (*id.*, p.127) and defines as a way of thinking of experiences that cannot transcend one's first-person perspective and that, consequently, affects their concept of such experiences (*ibid.*). We thus form *empirical* concepts by undergoing the experiences that they refer to, and this informs the distinctive first-personal character of these concepts. The concept of pain (and of mental states in general) is no exception to this. We acquire it by directly experiencing pain and therefore implement it through our first-person perspective. This consequently impacts the way we make sense of pain that we do not experience, namely, the pain of others – or, in McGinn's theory, of "third-person perspectives". Accordingly, since we cannot transcend our first-person perspective, we end up modelling another's experience of pain on our own experience of pain (*id.*, p. 129). We then conceive of a third-person perspective on pain by merely *extending* our first-personal concept of pain to others (*ibid.*). For instance, when somebody expresses pain in conversation, or in any other emotional exchange, we understand their being in pain through the features of it - its intensity, its unpleasantness, its temporality, and so on - that we have personally experienced. Thus, we understand others' pain, a third-personal kind of pain, through our empirical, thus inevitably first-personal familiarity with it. In doing so, we eventually treat our first-personal concept of pain as a univocal concept, equally applicable both to ourselves and to others. Yet, McGinn further asserts that this is "none too easy a thing to do" (Wittgenstein, 1953, §302), because we have a privileged access to our pain that we lack for the pain of others. As such, the way we understand others' pain shows to be different from the way we understand our own pain; we have, that is, "disjoint conditions" to assert our pain and that of others (McGinn, 1984, p. 136). This then challenges the alleged univocity of our first-personal concept of pain (*ibid.*); our concept of pain is thus to be classified as ambiguous (*ibid.*), in the sense that it cannot be unequivocally applied to both first- and third-person perspectives. This is what then leads McGinn to argue that a genuinely unequivocal concept of pain must be *neutral*, independent from both its different first- and third-personal character (*id.*, p. 137). To accomplish this, we must extrapolate the first-person and third-person conditions through which it is presented to us (*ibid.*). Yet, it seems that we cannot perform such an abstraction from the state of pain, for thinking of pain seems to always imply either a first-person or a third-person perspective (*ibid*). Indeed, I either think of pain as *my* pain – as pain that I directly undergo – or as another's pain - as pain that I do not, and cannot, undergo. Following

Wittgenstein (1953, §302), McGinn also immediately discharges the possibility of *imagining* another's pain, because imagination inherently contains the "inexpungible stamp" of our first-person perspective (Avramides, 2001, p. 227). We therefore seem to always think of pain as being either our own pain or someone else's, as inevitably involving either a first- or a third-person perspective (McGinn, 1984, p. 137). This leads to an impasse. On the one hand, combining the two perspectives in the same concept of pain *destroys* its (apparent) *univocity*, because of their two distinct conditions of verifiability. On the other hand, prescinding from them is highly counterintuitive (*ibid*). As such, our concept of pain tends to oscillate between either perspective (*ibid.*). McGinn thus concludes that a successful resolution of CP is unfeasible (*ibid.*).

2. Rejection of McGinn's Conceptual Ambiguity

In the previous section, I have clarified McGinn's interpretation of CP. I will now proceed to evaluate and reject it.

A helpful start for this evaluation is the ambiguity that McGinn alleges in our use of the concept of pain (henceforth referred to as "conceptual ambiguity"). As we have ascertained, this brings attention to the gulf between our understanding of our own pain and our understanding of the pain of others. We know our pain by directly feeling it, but we are not equally entitled to feel the pain of others. Yet, the source of our knowledge of others' pain remains an open question. McGinn himself never addresses it; still, he hints at a possible answer by mentioning that behaviourist accounts of CP imply an exclusively third-person perspective (McGinn, 1984, p.137). According to its standard reading, behaviourism purports that the ascription of psychological states demands corresponding behavioural evidence, but this does not hold conversely (Graham, 2023, §1). That is, behavioural evidence is necessary (and sufficient) for the ascription of mental states, whereas mental states are not necessary for the explanation of behavioural evidence. It is then fair to assume that McGinn considers such behavioural evidence to ground our understanding and subsequent ascription of pain to others. We can thus ascribe pain to others only if we have behavioural confirmation that they are in such states. In this sense, McGinn defines the first-person/third-person dichotomy as an inner/outer dichotomy: we know our pain through inner introspection and the pain of others through their behavioural manifestations – for instance, through their bodily movements or their utterances.

However, it is dubious whether such a behaviourist account ultimately supports McGinn's argument. For it seems to me untrue that behavioural evidence does not require the ascription

of mental states in order to count as such. Consider, for example, the stillness of a rock, or the flowing motion of a river. These constitute "outer evidence", much like another individual's bodily movements and utterances do. Nonetheless, in the case of the rock and the river, we do not (traditionally) classify such evidence as "behavioural" as we do in the case of the individual. I believe that this happens because we do not consider the rock nor the river to be sentient beings – namely, beings who appreciate some sort of mental life. Why we do not is an enquiry that I will broach later. For the time being, these contrasting examples intend to show that, to recognise another's behaviour *as* behaviour, we must already presuppose the exercise of psychological activity in them. In other words, we must already consider them as a *being,* as a subject to mental sensations and experiences. Consequently, not only is the ascription of mental states necessary for outer manifestations to count as "behavioural", but such ascription is also logically prior. We understand others' external manifestations as behaviour if we have already appointed them as a being. According to this interpretation, we would then employ others' behaviour to verify their experience of pain at a given time, rather than their aptitude for feeling pain *at all.* This consequently reduces McGinn's conceptual ambiguity to an instance of epistemological scepticism, that is, to merely doubting whether we know what mental states others are experiencing in a specific moment, and not our ability to unequivocally ascribe such states (i.e. pain) to them. As a result, its pertinence to the issue at hand inevitably appears undermined.

Such pertinence is then ultimately dismissed when we realise that to acknowledge others as beings amounts to attributing them a first-person perspective rather than a third-person one. To be a subject of experiences and mental states, a being must be indeed aware of them, thus think of people as their own, thus produce first-personal concepts of them. All of this, as we have seen, is attainable only through introspection and its joint first-person perspective. The ability to undergo experiences therefore relies on the appreciation of a first-person perspective. In this light, a being is, so to speak, a first-person perspective – a first-personal way of being subjected to experiences. Our reasoning thus resumes as follows: we can distinguish another's behaviour only if we already presume that they undergo mental states and experiences, that is, that they are a being. Yet, we can presume that they are a being only if we already recognise them as appreciating a first-person perspective. It then becomes evident that this first-personal character is the foundation for any attribution of mental states (and behaviour) to others. We must first acknowledge that others possess a first-person perspective, *just like we do.* We must recognise them, that is, as another *me.* And we do so by extending our first-person perspective to others. While this might resemble the practice that McGinn himself refutes, the two are substantially different. For his view entails the sole attribution of our first-personal concept of pain (and of mental states) to others, whereas

mine expects us to extend the very presence of a first-person perspective to others. And since we cannot transcend our own first-person perspective, we understand others' to parallel ours. That is, we understand others to have the same first-person perspective, to undergo experiences in the same way we do. If this is the case, it then follows that we are justified in modelling the pain of others on our own experience of it. We then come to know what it is for others to feel pain because we know what it is for us to feel pain and we view others as being *like us*. It reveals that we need not appeal to external evidence to understand others' pain, but, rather, that we understand both our pain and the pain of others through our first-personal concept of it, namely, through our introspection. As such, when we ascribe pain to ourselves and to others, we are employing a univocal first-personal concept of pain – first-personal indeed, but consistent, nonetheless. While we do understand that other people can undergo pain, we understand the *features* of their pain only via our own experience of it. For instance, we can make sense of somebody feeling pain in their twisted ankle, but we cannot make sense of the pain of a twisted ankle itself until we, too, suffer it. Or, if we have previously suffered it, then we can *only* understand it through the subjective features that such pain manifested with in our personal experience. If this is the case, then McGinn's already weakened conceptual ambiguity is to be systematically rejected, and his quest for a neutral concept of pain dismissed as ineffective.

3. Objections & Subjectivity: A Different Problem of Other Minds

I have just argued that first-person perspective warrants the attribution of mental states and behaviour to others, and that it induces an unequivocal use of our first-personal concept of pain. I will now present two connected criticisms against this thesis, through which I will aim to explore an alternative understanding of CP.

According to the thesis I have suggested in Section 2, we extend our first-person perspective to others and model their pain (and other experiences) on our own. I believe that this portrays a realistic picture of the way we usually generalise our way of relating to the world. We do it naturally and unavoidably because we do not know, *by ourselves*, of any other possible way to undergo such experiences. However, this conflicts with our intuition that there exist, indeed, various ways of undergoing experiences. A sharp example of this resides in the contemporary discourse between "typical" and "divergent" forms of neuro-processing, that is, between neuro-typicality and neuro-divergence – such as ADHD, Autism, Dyslexia, and so on. But how can we make sense of others' diverse first-person perspectives, if we are supposed to model them on our own? This then reflects the major criticism against my thesis:

that my thesis merely defers CP rather than truly resolving it. It translates it into an obstacle in comprehending the existence of first-person perspectives that are *another*'s, that are *like* ours but *not identical* to ours. In the case of the twisted ankle, for example, my thesis does not clarify how we are meant to understand other people's experience of this kind of pain as *different* from our own, if we can only make sense of such pain by directly undergoing it ourselves, thus rendering it *identical* to our own. In this sense, then, CP becomes an obstacle in comprehending the existence of first-person perspectives that are *another*'s, that are *like* ours but *not identical* to ours. In addition to this, my thesis also fails to explain how we can discriminate between sentient beings and non-sentient things – such as the rock and the river of our aforementioned example.

I believe that this objection raises a rather stimulating version of CP, which suspects our unequivocal use of our concept of first-person perspective. This is thus the version that I shall endeavour to answer now; for this matter, I will employ Davidson's discussion on empirical knowledge and belief as a starting point. In this respect, Davidson purports the inextricable connection between three kinds of empirical knowledge: knowledge of 1) one's own mind, 2) other minds, and 3) the world (Avramides, 2007, p.241). Applied to my own argument, these then can be interpreted as, respectively, knowledge of 1) subjectivity, 2) alter-subjectivity, and 3) objectivity. I will take these to mean 1) first-person perspective that is distinctively mine, 2) first-person perspective that is distinctively another's, 3) third-person perspective. These kinds are dependent on one another and mutually irreducible, so we cannot prescind from any of them (*ibid.*). He subsequently links knowledge and belief by interpreting the latter as a condition of knowledge (*ibid.*). However, for a belief to count as such, we must be able to recognise its contents; that is, we must understand what we and others believe (*ibid.*). Davidson holds that this is possible only through language and, particularly, through intersubjective (i.e. interpersonal) communication (*id.*, pp. 241-243). This is because of the reciprocal interpretation typical of this kind of setting (*id.*, p. 242). Indeed, in intersubjective communication, we contemporarily adopt the role of the speaker and that of the listener. As listeners, we actively interpret both the other's beliefs, which we recognise as being logically consistent with one another, and their meanings, which we understand as being shared responses to a shared world. As speakers, we witness the other operating these same processes to us (*ibid.*). For Davidson, this then results in our appreciation of all three kinds of empirical knowledge. That is to say, we come to know both our mind and the mind of others, while also acknowledging that we live in the same reality (*ibid.*). This is equally what grants our knowledge of subjectivity, alter-subjectivities, and objectivity. When we engage in intersubjective dialogues, in fact, we come to understand that our first-person perspective and that of others are not identical. They might resemble each

other in some respects, but then they are different in regard to others. We thus end up appreciating not only that others' first-person perspective is distinctively theirs, but also that our own first-person perspective is distinctively ours. That is, we end up appreciating their nature as, respectively, alter-subjectivity and subjectivity. Moreover, we also come to realise that we receive common stimuli from the world, to which we develop our own responses. As such, we come to interpret the world as objective, as preserving the essential features that we then experience. Rocks and rivers, we understand, partake in this objectivity, and thus contribute to shape our external reality.

This can be one response to the present understanding of CP. Nevertheless, one might still argue that such a response issues a self-refuting dilemma in my thesis. Indeed, admitting that we extend our first-person perspective to others seems to contradict my later assertion that we come to understand our subjectivity through interpersonal dialogues. However, favouring one over the other results in two counterintuitive scenarios. For, if we accept the former, our question on distinct first-person perspectives remains unsolved; if we accept the latter, we fail to elucidate why we generalise our first-person perspective. Either way, my thesis results as inconclusive and it should therefore be dismissed.

I believe that this objection confuses the terms "first-person perspective" and "subjectivity"; however, in my thesis I employed them with different meanings. In particular, the former is connected to our introspection and relates to our way of thinking of our experiences. We thus already appreciate it by ourselves, without needing to interact with others. The latter, instead, refers to the understanding of our first-person perspective as *distinctively ours*, which we can achieve only through intersubjective communication. This combination is not mutually exclusive. Indeed, while our only way of understanding other's pain is precisely through our own, because we do not have experience of any other pain than ours, this does not rule out our awareness that, despite our understanding, others might feel pain in ways that are very different from ours; and this awareness originates precisely from our interpersonal relationships, from intersubjectivity. Accordingly, instead of discarding my thesis, this objection comes to strengthen one of its core aspects, the coexistence of our univocal concept of first-person perspective and intersubjectivity.

Conclusion

In this essay I have rejected McGinn's interpretation of CP, particularly his verification principle. Instead, I have endorsed that, in extending our first-person perspective to others, we warrant an unequivocal attribution of our first-personal concept of pain to others. I have

then considered two objections against my thesis, both reliant on an interpretation of CP different from McGinn's and Wittgenstein's. The first one enabled me to implement my thesis with the notions of subjectivity, alter-subjectivity, and objectivity, thus explaining how we can make sense of both first-person perspectives distinct from our own, and of objectivity. The second one allowed me to clarify the distinction between "first-person perspective" and "subjectivity".

I therefore trust that my thesis has offered alternative insights on the conceptual problem of other minds, showing how it stems from our isolative experiences and then dies when we share them with others.

Bibliography:

Avramides, A. (2001) *Other Minds*. London, UK: Routledge.

Gomes, A. (2011) 'Is There A Problem of Other Minds?'. *Proceedings of the Aristotelian Society* 111: 353-373.

Graham, G. (2023) 'Behaviorism'. In *The Stanford Encyclopedia of Philosophy*, E. Zalta (ed.), Accessed at: https://plato.stanford.edu/entries/behaviorism/.

McGinn, C. (1984) 'What Is The Problem of Other Minds?'. *Proceedings of the Aristotelian Society* 58: 119-37.

Wittgenstein, L. J. J. 1953. *Philosophical Investigations,* G. E. M. Anscombe (transl.). New York, USA: Wiley-Blackwell.

Taking Linguistic Solipsism to its Limit

By Nina Stremersch

More than 100 years on, Wittgenstein's *Tractatus Logico-Philosophicus* continues to be a central text to philosophical inquiry. Stremersch brings a creative approach to this classic and cuts through the inaccessibility of the work in the process. While the *Tractatus* is typically associated with the logical positivists, who accorded the book a quasi-biblical importance, Stremersch introduces it to a very different branch of philosophy. Rather than taking early Wittgenstein's understanding of language as given, she introduces a phenomenological perspective more consistent with his later work and that of thinkers such as Husserl and Merleau-Ponty. In doing so, she creates a fresh perspective on the well-known solipsistic dilemma at the heart of the *Tractatus*.

In this essay I would like to ask the question, '*where do we end up, if we strictly follow out the implications of Wittgenstein's linguistic solipsism in the Tractatus?'* This inquiry leads me to argue that Wittgenstein's linguistic solipsism crucially determines his showing/saying distinction. My challenge to the distinction is then that it relies on a language which is publicly unintelligible, according to a phenomenological account of private language.

First, I justify my exegetical approach, which supports an anti-philosophical reading of the *Tractatus*, and expound the showing/saying distinction in the relevant detail. I then show that the distinction is crucially determined by the Tractarian Self: 'the Self' draws a limit to sense, or to what can be said. However, for Wittgenstein, this Self is the Self of linguistic solipsism, which coincides with pure realism (for what can be said). Subsequently I make the case that the solipsist/pure-realist account of language on which the showing/saying distinction then relies is unintelligible: on a phenomenological interpretation of private language, a meaningful language is necessarily intersubjective and public.

To begin, a remark on the relation of the *Tractatus* to the question of truth, which allows me to justify my exegetical approach. It is significant that many of the major schools of Tractarian interpretation systematise the work – and, what is more, systematically read the *Tractatus* in terms of *the question of truth*. For example, Hacker and Anscombe's reading of the *Tractatus* as pointing to ineffable *truths* (Anscombe, 2005; Hacker, 1972). Or the Logical positivists' exaltation of the Tractatus as securing the *truth* conditions for propositions (Ayer, 1972 p. 2).

In 1929, eight years after the publication of the *Tractatus*, Wittgenstein remarked that his "method [of philosophizing] consists essentially of leaving the question of *truth* behind and asking about *sense* instead" (Wittgenstein, 2006, pp. 105, 46). We do not yet hereby have the licence to ascribe this philosophic method to the *Tractatus*, but this question pushes us to ask more closely about the work's method and intent in relation to the problem of truth versus sense. The early Wittgenstein wrote to Russell emphasising that "the main point" of the Tractatus "is the theory of what can be expressed (gesagt) by prop[osition]s – i.e. by language – (and, which comes to the same, what can be thought) and what can not be expressed by prop[osition]s, but only shown (gezeigt)" (Wittgenstein, 1995, p. 98). This he believed to be "the cardinal problem of philosophy"– and this cardinal problem, expressed by the showing/saying distinction is, ultimately, a question of sense (Wittgenstein, 1995, p.98).

For now, we can understand 'sense' as the possible agreement of a proposition with reality (i.e. truth-aptness), and 'truth' as the actual agreement of a proposition with reality. Indeed, if we read the *Tractatus* as an anti-philosophic act, it consists not simply in denying truth to certain questions and answers of philosophy, like the positivists. Rather, it achieves an *inversion* of the concerns of philosophy itself. There is a trivialisation of *truth* over *sense* (Badiou, 2011, p. 116). The question is no longer 'What can we say that agrees with reality, and how can we demonstrate that agreement?', but rather 'What can we say about reality at all, and what right do we have to it'?. The *Tractatus* is then concerned with the question of truth to the extent that it takes philosophy's attempts to find truth, to a point of failure. Wittgenstein demonstrates the limit of such a truth-seeking method, or of *his* truth-seeking method, in according sense to metaphysics and the problems of philosophy. This anti-philosophic inversion also coincides with hostility to the dogmatism of systemisation (*ibid.*), since a systematic method must be optimistic about what it can find, which should reveal itself in the capacity to secure truth, and for that purpose must presuppose its own aptness to truth. For this reason, I should like to leave behind those systematising and systematically *truth*-oriented lines of interpretation, and so push Wittgenstein's project toward the limits of sense.

I will now provide an overview of the Tractarian showing/saying distinction, which draws a limit to sense, in the relevant detail for my subsequent argument. According to the *Tractatus*, the language is nothing other than the totality of propositions (4.001). Hence, what is ascribed to propositions is at the same time ascribed to language. The *Tractatus* understands propositions in terms of a picture theory (2.1–4.128). A picture is a model of reality (2.12), it *depicts* reality, and this is how propositions can express reality. A proposition has *sense* (i.e. is a thought) if it can correctly (i.e. truthfully) or incorrectly (i.e. falsely) depict reality (4). A proposition that cannot be depicted is *nonsense* (TLP Pref.)– it is not truth-apt, the concepts

of truth and falsity do not apply at all because there is no possibility that what it expresses exists.

Propositions are able to depict reality because the picture has something in common with reality (2.17). What they have in common is 'logical form' (4.12). Wittgenstein's key insight here is that the logical form of a picture itself cannot be expressed by a proposition (4.121). If we attempt to express the logical form of a picture propositionally, i.e. *say* it, we end up just repeating a proposition, rather than getting to its logical form. The logical form of a proposition cannot be said, because that would require a proposition outside of the world. But, since the world is constituted by the totality of facts depicted by propositions, propositions can only be in/of the world (1.1). What lies outside the world is then nonsense because it cannot be depicted. Therefore, the logical form of propositions cannot be said, only shown. The disjunction is thus drawn:

4.1212. "What *can* be shown, *cannot* be said."

What can only be shown, for Wittgenstein, is not restricted to logical form (Glock, 1996). It is just with this proposition that the distinction is first and most explicitly presented. Other propositions which are nonsense and can only be shown include pictorial form (2.172), the mystical (6.522), … and *solipsism* (5.62). It is the 'unsayable' of solipsism, which I now pass to consider as a primary ground of the showing/saying distinction.

Wittgenstein's affirmation of solipsism, I contend, determines the showing/saying distinction: 'the Self' of linguistic solipsism draws the limit to what can be said. This is manifest in the 5.6-5.4 discussion on 'the Self'. Here Wittgenstein rejects the Cartesian conflation of the representing subject with the psychological self (5.631). For him, "there is no such thing as the subject that thinks or entertains ideas", no psychological self (5.631). Similar to Kant, Wittgenstein then distinguishes between a metaphysical subject [transcendental for Kant] and a psychological subject [empirical for Kant] (Kant & Smith, 1934). It is the metaphysical subject, which is the limit of the world, which cannot be in the world (5.632). Like Kant's transcendental subject, Wittgenstein's metaphysical subject serves as the necessary structure for the possibility of speaking of the world, but for that same reason it cannot be said, i.e. cannot be in the world.

How, then, does Wittgenstein bring the self into philosophy, if this philosophical self is not part of the world, and "nothing in [the world] allows [us] to infer that it is seen by an [I]" (5.633)? Simply by "the fact that 'the world is my world'" (5.641). Indeed, we often claim to find in the world that which we forget we ourselves bring to the world. It is this forgetfulness

which leads us to claim certain truths about the world, without being critical of what made it possible to access such knowledge in the first place. For Wittgenstein here, the solution is to say that the world somehow belongs to him. In this way he can maintain that his *Tractarian* language is the language which exhausts what can be said about reality. This recognition of the structural import of our subjective constitution, given by language, to what we say about the world, underlies Wittgenstein's following declaration.

5.6. "The limits of my language mean the limits of my world."

In the *Tractatus*, solipsism takes a linguistic turn. Linguistic solipsism consists entirely in the identification of *my* language with *the* language. The solipsist thinks he is the one; the linguistic solipsist thinks his language is the one– the one language constituting reality, or in context of the Tractatus, the one language which mirrors the logical form of the world. This collapse of intersubjectivity into a pure subjectivity is analogous to the movement of Tolstoy's *The Gospel in Brief*, a work dear to Wittgenstein. Namely, a collapse of *The Gospels* into *The (Tolstoy's) Gospel* (Tolstoy, 1961). The corollary, per the 5.6 identification of the limits of my language with the limit of my world, is that "[t]he world is *my* world" (5.62). It follows that what can be said is in the world; and that what cannot be said, only shown, lies outside the world. Therefore, that which is *in* my world has sense, can be said; what lies *outside* my world is nonsense, cannot be said, only shown. Thus the Self draws a limit, through *my*/*the* language, to sense.

Tractarian solipsism does not merely draw a limit to sense through language. Through its corollary, pure realism, it also determines the limit: "it can be seen that solipsism, when its implications are followed out strictly, coincides with pure realism. The self of solipsism shrinks to a point without extension" (5.64). If *my* language is *the* language, and I am the limit of the world with which this language is identified, then all propositions are independent of a unique Self. Pure realism obtains, in that propositions are dependent on *my* language, but since *my* language is *the* language, those same propositions are nonetheless independent of *me*. Thus, for example, "I have pain." becomes synonymous with 'There is pain.' (Wittgenstein, 1969, p. 67). As no unique ownership subsists, reference to an *I* is superfluous (id., p. 67). Values, such as those of "ethics/aesthetics", which are subjectively constituted, also become nonsense (6.42-6.421); they cannot lie in the world, for 'the Self' is not itself within the world (5.632). Hence, through that same movement whereby the *Tractatus* casts 'the Self' out of the world, so too is the sense of the world, or the mystical, cast out of the world (6.41), for "I am my world" (5.63). Thus, the *Tractatus'* affirmation of solipsism crucially draws a limit to sense, and thereby determines the (*my*) showing/saying distinction.

Given this, I now challenge the showing/saying distinction by making the case that the solipsist/pure-realist language on which it operates is unintelligible. My argument consists in a phenomenological turn on the private language arguments levied by the later Wittgenstein.

The language of the linguistic solipsist is a private language: it is "that language which I alone understand" (5.62). The *Tractatus* claims that such a language has sense, even though it is intelligible only to *me*, because it consists in mirroring the absolute logical structure of the world, which is thus ostensibly universally intelligible.

Yet a meaningful language fails to obtain. Consider Husserl's claim that all meaning is intentional, i.e. is *about* or *directed toward* something (Husserl, 1970, I I §1-12). On this argument, language does not have meaning, or sense, merely by *mirroring* the logical structure of the world. Rather, the sense of language is affected through the publicly oriented act of speech (Merleau-Ponty, 1962, p. 206). When I utter a word, say 'tractor', I have the expectation of expressing something which is somehow already intelligible outside of my mind, i.e. public. I express something which is shared, as I suppose to myself possible interlocutors who understand a personal yet common meaning of 'tractor'. In that sense it is then irrelevant to the meaning of the word that *I* am its articulator. A meaningful language is therefore about something shared and public. It is this sense that "speech is an institution" (Merleau-Ponty, 1962, p. 213).

Thus, an intelligible language has an assumption of sharedness, communality – an oversight of the *Tractatus* only realised in the later *Philosophical Investigations*, where meaning was taken as use, and use as public (Wittgenstein, 1958). The solipsist identification of subjectivity with objectivity fails to achieve publicity, for that objectivity is still not intersubjectively constituted. *I* am still restricted to a singular, thus private, viewpoint. Consequently, the universalisation of a private language itself remains private. The linguistic solipsist has merely identified what is intelligible to him with what is absolutely intelligible, but this fails to pass into public intelligibility – and thus the solipsist/pure-realist language falls short of meaning.

Consequently, given my demonstration of the showing/saying distinction's reliance on Tractarian solipsism, the distinction then rests on and is articulated through an unintelligible language. To that extent the nature of the distinction, being drawn in *my* language, is itself questioned.

This essay has then followed through on Wittgenstein's linguistic solipsism in the *Tractatus*. It has challenged the tenability of the Tractarian showing/saying distinction by demonstrating the primacy of linguistic solipsism in the determination of the distinction, and consequently arguing that the solipsist/pure-realist language on which the distinction relies is unintelligible. This is because a meaningful language is necessarily public, yet the solipsist/pure-realist language is private. A distinction which is determined by and articulated in an unintelligible language, has questionable tenability. I have thus marvelled not at *how*, but at *that*, the showing/saying distinction is drawn (6.44).

Bibliography:

Anscombe, G.E.M. (2005) *An Introduction to Wittgenstein's Tractatus*. South Bend, in: St. Augustine's Press.

Ayer, A.J. (1972) *Language, Truth and Logic*. Middlesex: Penguin Books.

Badiou, A. and Bosteels, B. (2011) *Wittgenstein's Antiphilosophy*. London: Verso.

Glock, H.-J. (1996) *A Wittgenstein Dictionary*. Oxford, OX, UK: Blackwell Reference.

Hacker, P.M.S. (1972) *Insight and Illusion: Wittgenstein on philosophy and the metaphysics of experience*. Oxford, Clarendon Press.

Husserl, E. (1970) *Logical Investigations* J. N. Findlay *Vol I*. Translated by J.N. Findlay. London and New York: Routledge.

Kant, I. (1934) 'Transcendental Deduction', *Critique of Pure Reason* Norman Kemp Smith. Translated by N.K. Smith. London: Macmillan.

Merleau-Ponty, M. (1962) *Phenomenology of Perception* Colin Smith. Translated by C. Smith. London and New York: Routledge & Keegan Paul.

Tolstoy, L. (1961) *A confession ; The Gospel in Brief ; and What I Believe* Aylmer Maude. Translated by A. Maude. London: Oxford University Press.

Wittgenstein, L. *et al.* (2006) *Culture and Value*. Malden etc.: Blackwell Publishing.

Wittgenstein, L. (1958) *Philosophical Investigations* G. E. M. Anscombe. Translated by G.E.M. Anscombe. Basil Blackwell Ltd.

Wittgenstein, L. (1969) *The Blue and Brown Books* G. E. M. Anscombe. Translated by G.E.M. Anscombe. Oxford: Blackwell Publishers Ltd.

Wittgenstein, L. (1998) *Tractatus Logico-Philosophicus* D.F. Pears and B.F. McGuinness. Translated by D.F. Pears and B.F. McGuinness. London: Routledge & Keegan Paul.

Wittgenstein, L. (1995) *Wittgenstein in Cambridge: Letters and Documents 1911-1951* Brian McGuinness. Translated by B. McGuinness. Oxford: Blackwell Publishing.

N.B. all purely numerical citations, e.g. 5.31, refer to propositions of the *Tractatus Logico-Philosophicus* (Wittgenstein 1998).

On the Schematism of the Pure Concepts of the Understanding: An Expository Essay

By Miriam Treitinger

What does it mean for one event to follow another? By what means do we conclude that, to use Treitinger's example, pulling a gun's trigger *causes* a gun to fire? In this essay, Treitinger considers Kant's solution to this problem: the faculty of imagination. Though perhaps arcane, the *Transcendental Schematism of the Pure Concepts of the Understanding* is a key passage in Kant's revolutionary text, *The Critique of Pure Reason.* Lucidly written, Treitinger's essay is an exciting accessible exposition and criticism of a salient section of Kant's difficult text.

Kant tells us that the schematism of the understanding is hidden deep in our soul (Kant, 1998 [1787], pp. B180–181/A141). It is a product of the imagination that allows us to connect the sensual to the conceptual. Yet, the details of its origin remain unclear. The following is an exposition of the section of Immanuel Kant's *Critique of Pure Reason* that is generally (and synonymously to the idea it is presenting) known as the schematism and constitutes the first chapter of the *Analytic of Principles*. After introducing the question Kant tries to answer in the schematism—that of how exactly we utilise our imagination to apply categories to objects—I will provide some context necessary for understanding this part of the book. The latter will consist of one paragraph on the cognitive faculty of the power of judgement, which relies on the schematism, and one on homogeneity through which schemata are connected to both intuitions and concepts. Subsequently, I will explicate the problem the schematism tries to solve before explaining the solution Kant offers in detail. Then, hoping to facilitate a more nuanced understanding of the schematism, I will investigate and argue in favour of Gardner and others' criticism that is targeted at the schematism possibly undermining the concept/intuition divide.

The *Deduction,* which precedes the *Analytic of Principles,* has shown that we do in fact apply categories (or non-sensible, pure concepts of the understanding [as outlined in the *Doctrine of Elements*]) to objects given to us in intuition. Yet, despite all the efforts of the project of the *Deduction* to explore the relation of the categories to objects of empirical intuition, at

least one question remains: What exactly is the mental act by which the imagination connects the categories to objects? In the schematism, Kant provides a controversial answer to this, which has often been deemed unsatisfactory.

Before analysing the schematism, I will, in a similar fashion as Kant himself in the introduction to the *Analytic of Principles*, outline the role of the power of judgement in transcendental logic, as it contributes substantially to the context in which the schematism is presented. For Kant, the power of judgement is one of the three faculties of cognition (which are the understanding, the power of judgement, and reason). As it is concerned with determining whether a concept applies to an object, the power of judgement only plays a role in transcendental logic and is not relevant in general logic.[1] The power of judgement allows us to determine whether a given rule or concept applies to a specific object or not. Kant provides the example of a judge, which I want to lean on to illustrate this (B173–174/A134): A judge knows the definition of murder. This would be analogous to a concept. Let us suppose the judge is now presented with a specific case in which a murder has been suspected and is shown photos of the crime scene, hears the suspect, etc. This is analogous to an object. Now, the judge has to decide whether the definition applies to what she has intuited (seen, heard, etc.) of the scene. She has to make a judgement. This judgement can adhere to reality or not. The power of judgement is the ability to correctly assess whether the concept applies to the object. One might point out that the example of the judge fails to represent the automatism and spontaneity of the power of judgement, which operates in an almost unconscious way. For those who struggle to conceive of such a cognitive faculty without any aid, I suggest imagining a very small judge, working in the doorway between the conscious and the unconscious, who works very, very quickly—in fact so quickly that the judgment is made in the exact same moment in which the evidence is first observed. We cannot even really talk about speed here, as no time is passing without conceiving of the object in terms of the judgment, as judgment is one of the conditions underlying our spatio-temporal perception.

The second idea that is crucial to explore before explaining the schematism is that of homogeneity. It is important to note that Kant uses the word in varying ways in the *Critique of Pure Reason*. I will only explore the more colloquial sense of the term as used in the schematism. While it is not always easy to make a correct judgement when it comes to the

[1] Transcendental logic is concerned with how thought relates to objects, while Kant uses the term 'general logic' for what is commonly just referred to as 'logic' or 'formal logic' (Gardner, 1999, p. 126).

application of empirical concepts to intuitions, on a surface level, the way in which we arrive at such a judgement is rather straightforward. This is because all representations of an object to which a concept applies are already contained in that concept (B176/A137). This is what Kant, in this section of the book, refers to as homogeneity. In order to judge whether an empirical concept applies to an object, we only need to be sufficiently aware of the concept's definition and hope that our senses provide us with the necessary information to determine whether the definition applies. To reiterate: the representation of the object, which is its geometrical form, is already part of the concept.

Kant now tells us that the categories (e.g., causality or possibility), on the other hand, are entirely unhomogeneous, i.e. that no representations are subsumed under them (*ibid.*). But, as we already know, we do in fact somehow apply the categories to representations or objects. The question is: *How*? If we take the statement 'the gun fired because I pulled the trigger', we might somehow believe that we have *observed* causality. This idea, however, is rather flawed. Pulling a trigger on a gun and the gun subsequently firing are two processes that have been observed separately. To say that the first caused the latter is nothing that can be reasonably concluded exclusively from what has been intuited here. One cannot observe or intuit causality or 'becauseness'. The statement 'the gun fired because I pulled the trigger' is of course reasonably assumed to be true, but the mental processes underlying such a statement have to be more complex than the mere evaluation of whether a representation is contained in the category of causality.

The answer to the question 'what did the mind do in order to (correctly) conclude that pulling the trigger caused the gun to fire?' may serve as an illustration of the answer to the general question of how exactly we apply categories to objects. During the following explanation of the idea of the transcendental schemata, I will return to the example of the firing gun. In the analysis of the schematism, I will use this example in two ways: First, I will use it to illustrate the schematism, and second, I will use the schematism to give an interpretation of what would have happened here according to Kant. I chose this example because it seems rather easy to grasp and work with. However, we have to bear in mind that this example refers to an empirical intuition; and while the schematism applies to both the *a priori* and the *a posteriori*, it is foremost investigated with the *a priori* in mind.

Kant argues that, in order to apply categories, which are *a priori* and non-sensible, to objects, which are sensible, we make use of a so far unmentioned method of mediation between the two—the application of transcendental schemata (B177–178/A138–139).[2] To

[2] The terms 'schema' and 'transcendental schema' are synonymous. 'Schemata' is the plural of 'schema'.

fulfil this role, schemata have to be both *a priori* and sensible in order to be homogeneous with both the category they correspond to and at least some aspects of the objects to which they apply. Similarly to concepts, we should think of schemata as necessary and universal rules.

Transcendental schemata are the result of, insofar as possible, conceiving of the categories in terms of the necessary conditions for the possibility of experience, i.e., space and time. But the categories cannot possibly be conceived of in terms of space, as space only determines the conditions for outer intuition to which the categories, as pure concepts of the understanding, have no connection.[3] Therefore, categories can only be conceived of in an intuitable way if they are conceived of in terms of time. This is because time underlies all (i.e., both inner and outer) intuition. This is why transcendental schemata may also be called 'transcendental time-determinations'. In the chapter, Kant lists the schemata for all concepts. For example, the schema corresponding to the category of negation is 'non-being' (in time); therefore, the absence of sensation. The schema corresponding to the category of plurality (and indeed to all categories of quantity) is 'number'.

As representations of categories, schemata ought not to be confused with images. While images are pictorial representations, schemata are non-pictorial representations that cannot exist anywhere but in thought. They are prior to images, and, in turn, images presuppose schemata. Schemata represent concepts adequately, while images are inadequate representations of concepts.

If we take the pure geometrical concept 'triangle' and then encounter an image of a triangle, this image is not an adequate representation of the concept. In the case of the image of the triangle given to us a posteriori (e.g., in the form of a drawing), we might conclude that some inaccuracies resulted from the mechanical process of producing the drawing (e.g., the lines might not be perfectly straight). But even a mental image (and those are the kinds of images Kant is mostly concerned with) of a triangle that naturally avoids all of these kinds of error can only ever be one instance of a triangle and can therefore not be representative of the concept 'triangle', which subsumes all possible triangles. The image fails to represent the universality of the concept. We still know the concept 'triangle'—but while we can visualise an infinite number of possible triangles, none of these images can adequately represent the concept. The accurate representation would be the schema correlating to the

[3] Kant explains the distinction between inner and outer intuition as well as their relation to space and time in the *Transcendental Aesthetic*.

concept 'triangle'. The schema is also what we apply to the image to determine whether it can be connected to the concept or not. It might be relevant to note here that Kant himself uses the example of the concept of a triangle and thereby confirms something that within this framework seems necessary—that schemata do not only allow us to apply categories to objects, but that they further also allow us to apply the pure geometrical concepts to representations (B180/A141). We can think of those schemata that correlate to the pure geometrical concepts as arrangements of the schemata correlating to the mathematical categories (those of quantity and quality).

Let us return to the example of the firing gun. What is clearly (according to Kant) happening here is that we are applying the category of causality to the object. As we now know, we are doing this by using a schema that is homogeneous with both the category and the object. The schema, so to say the temporal interpretation for causality is "the real upon which, whenever it is posited, something else always follows," or in simpler terms: whenever there is A, B always follows (B183/A144). So, if indeed every time I pull the trigger the gun fires, the power of judgement allows me to determine that the concept applies. Therefore, I would be justified in saying that the gun fired *because* I pulled the trigger, rather than merely that one thing happened after another thing happened.

One might associate talk about causality with induction and David Hume. Asserting causality in the empirical world always relies on induction. Thus, the problem of induction applies (Morris and Brown, 2001). Just because the gun has always fired when I pulled the trigger in the past, it does not necessarily have to do the same in the future. And in fact, from what we know about guns, it is very easy to conceive of an alternative scenario, as perhaps I forgot to load the gun before pulling the trigger. But Kant is not particularly concerned with this, for his aim is to figure out the rules underlying judgement. He is not giving instructions on how to make accurate judgements about guns or anything else in the world. All that is implied by the schema of causation is that *if* the gun does really always fire when I pull the trigger, (that is, in the past, the future, and indeed at any time I do this in any possible world) then we can correctly judge that the gun fired *because* I pulled the trigger.

In the following, I will present and argue in favour of the view that the schematism undermines the concept/intuition divide fundamental to the *Critique of Pure Reason*.
Sebastian Gardner notes that, on a superficial level, Kant's proposal is quite straightforward, yet, on the other hand, puzzling when considered more carefully (Gardner, 1999, p. 170). He asks whether a transcendental schema is a) a thought about time or rather b) time as thought in a certain way *(ibid.)*. As a 'special' kind of *a priori* thought, transcendental schemata are

supposed to mediate between concepts and representations, yet either option seems to assimilate it to one or the other side of the division *(ibid.)*. Option a) would assimilate it to the side of concepts, option b) to that of representations. To that, Kant might say: a transcendental schema, when applied to an intuition, is a thought about the existence, non-existence, or possibility of existence of one or more things in the world in and/or through time. It is both sensible and intellectual, and therefore neither of the above options apply. Gardner and others interpret this as Kant saying that transcendental schemata are neither concepts nor representations, but rather a third thing. Therefore, Kant would imply that his original division is not exhaustive (*ibid.*).

In response to this criticism, one might now mistakenly think the following: 'Schemata, while not categories, are in fact concepts, which, like empirical concepts, do apply to things given to us in intuition. They differ from categories insofar as they are sensible, and they differ from empirical concepts insofar as they are pure. It would therefore seem appropriate to say that schemata are another kind of concept, besides the categories and empirical concepts, but not that they are a third thing besides concepts and representations. Therefore, transcendental schemata do in fact not challenge the concept/intuition divide.'

At first, this might seem rather convincing. However, this view results from a misunderstanding of the way in which we apply empirical concepts to sense-input. As pointed out above, the representations of an object are homogeneous with the concept under which the object is subsumed. It is not the object in itself, but rather its geometrical forms as represented to us, that are homogeneous with the concept. As the pure geometrical concepts require schemata to be applied to representations, and we in turn only make sense of the geometrical forms of objects through schemata, they are also essential for the application of empirical concepts to representations and in fact for the formation of empirical concepts in the first place. Transcendental schemata therefore must indeed be a third thing besides concepts and representations, as without them no kind of concept can be applied to any intuition. The implications of this conclusion have been argued to result in several fundamental problems with Kant's transcendental philosophy, to which the concept/intuition divide is essential.

Another point of criticism that I will only briefly mention here is the schematism's comparative underdevelopment. While Kant provides accounts of the derivation of both concepts and intuitions, he does not do the same for the schemata (*ibid.*). The transcendental schemata seem to result somehow from both intuition and concepts, yet their exact origin remains unexplored in the *Critique*. Kant himself states: "This schematism of our

understanding with regard to objects and their mere form is a hidden art in the depths of the human soul, whose true operations we can divine from nature and lay unveiled before our eyes only with difficulty" (B180–181/A141). And as we are now (hopefully) aware of the importance of the schematism, we can see how Kant's entire transcendental philosophy might suffer from the consequences of this neglect.

To conclude, after spending most of this essay presenting the schematism, I critically evaluated its role and situation in Kant's transcendental philosophy. After the clarification of essential terms and a subsequent explication of the main contents of the relevant section in the Critique of Pure Reason, the objection of the schematism undermining the concept/intuition divide has been considered. I concluded that, although it might, at first glance, seem convincing that schemata are a kind of concept, this is ultimately unconvincing because concepts and representations cannot be connected by something that is itself either a concept or representation. The schematism does therefore in fact undermine the concept/intuition divide. Lastly, a note on the schemata's underdevelopment underscored the comparative neglect of this essential part of the philosophy of transcendental idealism as presented by Kant.

Bibliography:

Gardner, S. (1999). *Routledge Philosophy Guidebook to Kant and the* Critique of Pure Reason. Routledge.

Kant, I. (1998 [1787]). *Critique of Pure Reason*. Translated by P. Guyer and A. W. Wood. Cambridge University Press.

Morris, W. E. and Brown, C. R. (2001). 'David Hume', in Zalta and Nodelman (eds) *The Stanford Encyclopedia of Philosophy* (Winter 2023 Edition). Metaphysics Research Lab, Stanford University. Retrieved from https://plato.stanford.edu/entries/hume.

Caught Between Two Problems: Can We Salvage Ayer's Emotivism?

By Eoin Moran

Emotivism is a non-cognitivist theory that seems compelling because it acknowledges the human characteristics of desire and rational failure that tend to threaten morality. Moran presents us with two objections against emotivism in a novel way. While the theory may withstand each objection on its own, their combined impact ultimately makes it untenable. This approach is compelling and original because it reminds us that we can rethink problems by using objections as tools in creative ways.

Introduction

This paper will be a critical review of A.J. Ayer's emotivist school of metaethical non-cognitivism. Ayer held that moral judgements are not beliefs (hence 'non-cognitivist') that cannot be described in terms of truth or falsity but, instead, only mere expressions of attitudes or emotions.

Ayer's reasons for believing this to be true will be disregarded in this paper. The focus here will not be on his claims for emotivism over rival theories but on two objections directed towards his theory. Specifically, this paper will first examine the objection concerned with wishful thinking (henceforth, 'wishful thinking objection' or WTO). This objection raises an issue as to how Ayer's moral judgements, being that they are non-cognitive, may mesh with our actual cognitive beliefs, especially when such moral judgements are used as premises in argument alongside these beliefs. Furthermore, this paper will relate the WTO to another preponderant objection to emotivism: the Frege-Geach Problem (henceforth, FGO). This objection points to some deeper flaws in Ayer's work insofar as it allows for the changing of the meaning of a word or statement from one premise to the next in an argument.

I posit in this paper that when these two objections are kept separate, they are not hugely effective, but, together, they provide a devastating attack on emotivism, hence my choice to use both. This paper will begin by examining the WTO and whether emotivism may be salvaged from this objection alone. Following this, I will show that the WTO may, on its own, be overcome by the emotivist (read: Ayer). However, I will then proceed to demonstrate how our circumvention of the WTO befalls us to another objection, the FGO. I will explain the FGO and how the rejection of it only brings us back to the WTO. As such, having shown

that the emotivist cannot overcome the WTO without facing the FGO, nor bypass the FGO without thereby being subject to the WTO, this paper will conclude that emotivism becomes untenable once both objections are considered together.

Ayer's Wishful Thinking

Ayer's emotivism states that moral judgements are not like ordinary beliefs, such as 'the chair is beside the table,' and that they are, in fact, not beliefs at all. While 'the chair is beside the table' can be said to be either true or false upon examination of the facts at hand, moral judgements, like 'it is wrong to commit arson on the chair and table,' cannot, according to Ayer, be said to be true or false. Instead, they are non-cognitive judgements akin to feelings (Ayer, 1936, pp. 64-72; Miller, 2003, pp. 26-38). So really, when we say 'it is wrong to commit arson on the chair and table,' we are not speaking in terms of belief in any real sense, but of our own feelings with regard to seeing the chair and table set on fire, specifically here that we disapprove of it (Ayer, 1936, pp. 64-72).

However, Cian Dorr noted a problem in this framework. Dorr wanted to know how it could be the case that we have a seemingly valid argument in which one of the premises is what Ayer takes to be a non-cognitive, non-belief, while still being able to maintain the validity of the argument (Dorr, 2002). One way to present such arguments is in the form of *modus ponens*:

Modus Ponens (MP)-WTO

P1. If arson is wrong, then the police will arrest arsonists.

P2. Arson is wrong.

C. The police will arrest arsonists.

This line of reasoning is a valid form of argument, wherein the premises necessitate the conclusion, by virtue of its form as *modus ponens*. Dorr asks us to imagine a cognitivist account of the thought process behind accepting this argument: Suppose that P1, 'if arson is wrong, then the police will arrest arsonists,' is learned by person X. Then, after some time, person X comes to accept what we may take here to be a cognitive judgement in P2, that 'arson is wrong' (Dorr, 2002, p. 98). This, we would think because of the form of the argument, allows X to rationally accept C, that 'the police will arrest arsonists.'

However, Dorr thinks that this obvious, uncontroversially valid argument may only be so in a cognitivist framework, something that the emotivist cannot account for. Dorr shows that the emotivist (i.e. Ayer) must reject this undoubtedly valid argument in explaining a non-cognitivist account of such a thought process (Dorr, 2002, pp. 98-99). Suppose that person

Y has the belief of P1. Then, for whatever reason, Y begins to hold what the non-cognitivist would be committed to saying is the attitude/feeling of P2. Finally, Y accepts C, completing the argument (Dorr, 2002, pp. 98-99; Van Roojen, 2018). Dorr maintains that such a thought process is wholly irrational, given that the emotivist framework is accepted. What has happened is that Y went from believing a fact in P1 to having a certain feeling in P2, then taking P1 and P2 in conjunction to accept C (Dorr, 2002; Van Roojen, 2018). This is not the way logic ought to work. Dorr explicates:

*"*It is often rational to modify your views about one part of the world so that they cohere with your views about the rest of the world.*"* (Dorr, 2002, p. 99) (Person X – the cognitivist)

"It is irrational to modify your views about the world so that they cohere with your desires and feelings. That is wishful thinking." (Dorr, 2002, p. 99) (Person Y – the non-cognitivist)

Should the emotivist try to escape the WTO, they will face a problem with two horns. First, the emotivist could deny the validity of this *modus ponens*, flying in the face of what is accepted by almost all contemporary philosophers as a perfectly valid argument (Guam, 2014, p. 35). Alternatively, they must say that it *is* valid for person Y to accept MP-WTO, even if between P1 and P2 nothing new is learned, but only a feeling is engendered within Y (Guam, 2014, p. 35). This option bears heavy implications for the emotivist, as to make their theory work, they would have to adhere to the deeply problematic position that we *can* derive fully descriptive beliefs about the world, like C, from what they hold to be mere feelings of approval and disapproval. Thus, the choice for the emotivist is either to deny the validity of a perfectly valid line of reasoning, or to accept its validity and affirm that we can wishfully pretend that something is true simply because we have a feeling that it is true. Since neither option seems promising, Dorr suggests abandoning emotivism (Dorr, 2002; Guam, 2014; Van Roojen, 2018).

Escaping Wishful Thinking

It is now up to the emotivist to respond to the charges of the WTO. Having seen that accepting either horn of the dilemma makes the emotivist position untenable, what could the emotivist possibly do to recover their theory from this challenge? Well, one way in which Ayer, determined as ever to salvage his emotivism from the WTO, may respond is by stating that moral terms are not *always* used to assert actual emotions. For instance, in MP-WTO, Ayer may suggest the argument goes as follows:

MP-WTO*

P1*. If 'arson is wrong', then the police will arrest arsonists.[1]

P2*. 'Arson is wrong.'[2]

C*. The police will arrest arsonists.

This response appears prima facie to have resolved the emotivist's trouble with the WTO. Formerly it was the case in MP-WTO that person Y stumbled into believing the conclusion with only the first premise and a feeling (the second premise) in mind, something we take to be wholly irrational. However, in this reformulation, MP-WTO*, the first premise is not so much fully a belief held by Y as much as it is a piece of hypothetical reasoning, still a belief, yet a more abstract kind. Furthermore, P2* is no longer exactly just a feeling, but a hypothetical feeling assumed for the sake of argument and that which Y may or may not actually hold. In this instance of hypothetical reasoning, the thinking toward accepting C* is no longer wishful and, while the second premise may or may not express a genuine emotion on the part of person Y, it serves the function of bringing us to the conclusion validly in such a hypothetical argument. This strategy employed of clarifying the premises seems to help the emotivist circumvent the WTO: we can both accept the validity of MP-WTO and not have to affirm the validity of wishful thinking. Thus, the emotivist's troubles have ostensibly been resolved.

However, the emotivist's work is not complete. For, although we may disregard concerns about the WTO for the moment, through their response above they now face another problem: the Frege-Geach Objection (FGO). How exactly we arrived at the FGO in our response will be clear upon explaining the problem itself. First and foremost, the FGO notes a significant problem for Ayer's emotivism in that such a framework struggles to consider how moral propositions may be embedded in larger sentences (Van Roojen, 2018). It puts forward that, while the cognitivist may have no problem with a sentence like 'if setting fire to the chair and table is wrong, then breaking the chair and table with a hammer is wrong too,' the emotivist may struggle immensely to make sense of it, since the moral sentence, which the emotivist takes to be solely feelings, is embedded in a complex sentence with the logical structure of a conditional (Guam, 2014, pp. 31-32; Miller, 2003, pp. 40-42; Van Roojen, 2018). Things only get worse for emotivism from here, for when we insert this

[1] (if [hypothetical thing a. – 'arson - is disapproved of], then [hypothetical thing b. – 'the police will arrest arsonists'])

[2] (either read as 'I, for the purposes of this argument, feel disapproval toward arson' or 'arson is disapproved of')

example sentence into a full *modus ponens* argument, the difficulty of creating such a non-cognitive account to handle these complex sentences grows:

MP-FGO

P1. If committing arson on the chair and table is wrong, then breaking the chair and table with a hammer is wrong.

P2. Arson is wrong.

C. Breaking the chair and table with a hammer is wrong.

The problem noted in this example is that the emotivist may have trouble not only dealing with the problem of moral sentences embedded into larger, complex sentences, but furthermore, how moral sentences can occur within complex sentences, as in the argument above, while having the same meaning throughout (Van Roojen, 2018). To elaborate, the FGO sets the challenge that the emotivist cannot give an account of premise 1 and premise 2 in the above argument such that the meaning of the moral sentence 'arson is wrong' does not change between premises. In premise 1, 'arson is wrong' is used unasserted as an antecedent in a conditional sentence. By comparison, in premise 2, 'arson is wrong' is used on its own purely, it seems, to assert the wrongness of arson. Clearly, the moral sentences here in MP-FGO are being used in different ways. In our response above to the WTO, MP-WTO*, this difference in meaning persists. In MP-WTO*, P1* was taken to be a hypothetical piece of reasoning, yet still a conditional nonetheless, while P2*, again taken to be hypothetical, was still a hypothetical assertion of a feeling. As such, in our response to the WTO through reformulating the meanings of the premises in MP-WTO, we have nonetheless encountered the FGO.

This difference in meanings is highly problematic (Guam, 2014, pp. 31-32; Unwin, 1999, pp. 337-338). If we change the meaning of the terms in the middle of an argument, we are fallaciously falsely equivocating. Miller provides the following example (Miller, 2003, pp. 40-42):

MP-Miller

P1. My beer has a head on it.

P2. If something has a head on it, then it must have eyes and ears.

C. My beer must have eyes and ears.

It appears that the emotivist here, both in MP-FGO and MP-WTO*, is committing a similar false equivalence, using the same sentence from P1 to P2 despite that sentence

differing semantically between the premises (Miller, 2003, pp. 40-42). Ayer's theory then, as was the case with the WTO, is forced by the FGO to once again condemn the validity of modus ponens.

Back to Thinking Wishfully

Hence, we must ask: what could Ayer do to recover emotivism now from the FGO? It seems that the clearest course of action would be to deny that MP-FGO features a changing of the meanings of terms throughout and that both terms are being used in an asserted context (i.e. being used to express emotion). While this response does not exactly solve all the problems the FGO poses in terms of embedding moral sentences as elemental pieces of complex sentences, it does at least appear to resolve issues of validity surrounding MP-FGO above, allowing the emotivist to sidestep the FGO for a moment.

However, while this response may grant the emotivist temporary relief from the FGO, it places them in the same position where we started, for they must now contend with the WTO. The emotivist can try to resolve MP-FGO by claiming that the meanings do not change and remain asserted throughout. However, all that this will do is set themselves up to once more face the WTO, in which they have to explain how it can be the case that the premises of the argument contain nothing more than feeling while retaining the validity of MP-WTO. We are, more or less, back to where we started, only now both premises are assertions of feeling.

At this point, it may start to seem that Ayer's emotivism is giving us more trouble than it is worth. It has to contend with both the WTO and the FGO. Both attack emotivism from two different fronts with the same result: Ayer's theory is made untenable by its inability to account for the validity of certain modus ponens arguments once the moral propositions contained within them are understood as emotive. I have already shown that any attempt to escape from the WTO only lands the emotivist into problems related to the FGO. Furthermore, I have now demonstrated that an attempt to overcome the FGO forces the emotivist to grapple with the WTO once more. Once the WTO and FGO are taken in conjunction to form this pincer movement, what can Ayer do to recover his theory? One route open to Ayer is to accept that we cannot argue in logic using moral terms whatsoever. This response appears to stop both the WTO and the FGO dead in their tracks. Since these objections raise issue with emotive moral language not functioning well in logical argument, Ayer asserting that moral sentences cannot be used to form logical arguments *at all* would make these problems non-issues.

The problem with this objection is not that it does a poor job of incorporating an account of moral discourse into Ayer's metaethics, but that it rejects the possibility of such an account altogether. In affirming that moral language is wholly incompatible with the language of logic, and that such a language is not apt for argumentation, Ayer would leave us with a metaethical theory that cannot cope with the how we use moral language in everyday life. This is highly problematic for emotivism. We would want our metaethical theories to be able to explain moral language as we use it. Ayer, in saying that moral terminology is incompatible with logic, would not then be delivering us with a satisfactory theory at all. Instead, we would be presented with a view that overlooks the intricate ways in which we use moral language. As such, if the only way to escape from the grasp formed through a conjunction of the WTO and FGO is to leave emotivism a shell of a theory, utterly incapable of translation into the normative realm of ethics, then our best course, it seems, is to abandon emotivism.

Conclusion

This paper had one main thesis, that Ayer's emotivism is untenable. The method set out to achieve this result has been through arguing that a conjunction between two different arguments against emotivism, the Wishful-Thinking Objection and the Frege-Geach Objection, proves to create a deadly attack on Ayer's theory. The choice to use these two objections in such a way has been properly justified through the flow of this paper's critique. I have shown that attempts to refute the WTO force the emotivist to deal with problems of the FGO. Reinforcing this, I have shown that in their escape from the FGO, the emotivist comes, more or less, full circle, once more having to defend their theory from the WTO. In other words, while the emotivist could navigate these problems as separate issues, taken as a pair it has been demonstrated that one cannot escape one without being brought about face to the other. As such, I maintain that emotivism has been shown in this paper to be untenable once caught in a trap between two objections.

Bibliography:

Ayer, A.J. (1936) *Language Truth and Logic* 1st Ed. London: Victor Gollancz Ltd.

Dorr, C. (2002) 'Non-cognitivism and Wishful Thinking,' *Noûs* 36(1):97–103

Guam, C. (2014) 'The Wishful Thinking Problem for Non-cognitivism: Does It Really Make Sense?' *KRITERION – Journal of Philosophy* 28(1):30-44

Miller, A. (2003) *An Introduction to Contemporary Metaethics* 1st Ed. Cambridge: Polity Press & Blackwell Publishing Ltd.

Unwin, N. (1999) 'Quasi-Realism, Negation and the Frege-Geach Problem', *The Philosophical Quarterly* 49(196): 337–352

Van Roojen, Mark., 'Moral Cognitivism vs. Non-Cognitivism.' In Edward N. Zalta (ed.) The Stanford Encyclopedia of Philosophy (Fall 2018 Edition). Available at https://plato.stanford.edu/archives/fall2018/entries/moral-cognitivism/. (Accessed 10th September 2023).

Reassessing Cantorian Cardinality: Conceptual and Pragmatic Objections to 'Transfinite Numbers'

By Robert Finan

Imagine trying to group all of the natural numbers in one set. Now, try comparing this set of natural numbers to the set of all integers. The question, which Finan's paper examines, arises: how might these two sets be just as big as each other? Don't believe it? Try counting them all and comparing these two sets! If and when you grow tired of this, Finan's essay will surely offer reprieve from the inevitable existential crisis! Therein, Finan considers Cantor's conception of cardinality in relation to infinity in set theory. Finan explores how Cantor's ideas about infinite sets, while revolutionary, lead to some rather counterintuitive conclusions and inconsistencies. The essay concludes constructively by proposing a way to retain the usefulness of Cantor's insights without its conceptual entanglements.

In this essay, I will critically analyse Cantor's conception of what it is for two sets to have the same cardinal number, as well as his related conclusions regarding 'transfinite cardinals'. I will argue against Cantor's view, highlighting its counterintuitive implications, inconsistencies, and the unnecessary upheaval it introduces to mathematical prose. I will primarily consider the supposed usefulness of Cantor's conclusions, as well as briefly addressing the supposed enlightening nature of said conclusions. Ultimately, I will argue that we ought to preserve the functionality of Cantor's conclusions by proposing a variant that distances itself from current rhetoric about 'transfinite cardinal numbers', particularly the application of relations of size to 'transfinite numbers'.

I will begin by briefly defining my terms:
For brevity, I will not rigorously define the notion of 'set', nor will I discuss its various conceptions; instead, I will take the notion of set as given.[1]

Subset (⊆): A is a subset of B, iff, for all $a \in A$, $a \in B$.

Proper Subset (⊂): A is a proper subset of B, iff, (A is a subset of B) & ($A \neq B$) & ($A \neq \{\}$)

Superset: A is a superset of B, iff, for all $b \in B$, $b \in A$.

[1] For the purposes of this essay, one may think of a 'set' as a 'collection of distinct objects', although this is not a formal definition per se.

Function: A mapping between sets wherein each element in the input set determines a unique element in the output set.
Bijection (Bijective Function): A one to one mapping from one set to another set, i.e. a function that is injective & surjective.
Injective function: A function, $f: A \rightarrow B$, wherein for all $x_1, x_2 \in A$, $(f(x_1) = f(x_2)) \rightarrow (x_1 = x_2)$
Surjective function: A function, $f: A \rightarrow B$, wherein for all $y \in B$ there is some $x \in A$, such that $f(x) = y$.

In addition to these definitions, I will briefly note that for our purposes 'cardinal number' and 'cardinality' are merely concerned with how many elements are in a set, with no further implications. Furthermore, there is potential for some confusion regarding the terms 'infinite' and 'transfinite'. Throughout this essay, 'infinite' refers broadly to that which is not finite, such as 'transfinite numbers', potentially infinite series, etc. 'Transfinite' refers to the specific concept outlined by Cantor: distinct 'transfinite numbers' that denote the cardinality of infinite sets (Cantor, 1932, p. 102).[2] Note here that 'infinite' is a general concept, whereas individual 'transfinite numbers' are specific numbers or values (Parker, 2008, p. 5).

I will further preface this essay by outlining Galileo's paradox.[3] Briefly, Galileo's paradox can be explained as follows (Galileo, 1954):
From studying finite sets, we come to two conclusions that are conventionally understood to be principles of sets in general. The relevant principles in this case, as outlined by M. W. Parker, are (Parker, 2019, pp. 375-376):

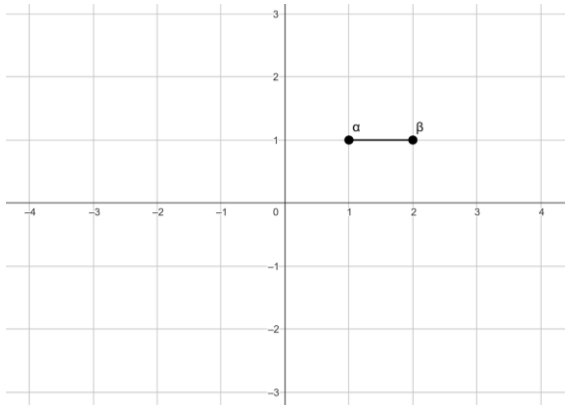
Bijection Principle (BP): If there is a bijection between A and B, then $|A| = |B|$.
Part-Whole Principle (PWP): If A is a proper subset of B, then $|A| < |B|$.

These principles lead Galileo to a paradox in the case of infinite sets. Specifically, Galileo draws attention to the set of natural numbers, N={1, 2, 3, 4, …}, and the set of square numbers, S={1, 4, 9, 16, …}. These sets can be put in bijection as $F: N \rightarrow S: f(x) = x^2$, and thus from BP, we find that $|N| = |S|$. Galileo then also notes that $S \subset N$, and thus from PWP, we find that $|N| > |S|$. It is here that we reach a contradiction that follows from BP,

[2] The cardinal numbers of infinite sets, such as the set of all natural numbers (if such cardinal numbers exist), are referred to by Cantor as 'transfinite'.
[3] I do this because Cantor—and subsequently Wittgenstein—engaged with Galileo's relevant works, as will become evident throughout this essay.

PWP and their application the relevant infinite sets, namely $(|N| = |S|)$ & $(|N| > |S|)$. Galileo understood BP and PWP to be fundamental and important; he prioritised these principles, seeing no need to have any conception of distinct infinite cardinalities. Thus, he concluded that we cannot apply relations of size to infinite sets (i.e. we cannot say there are *more*, *fewer*, or *equally many*, natural numbers and square numbers, we can only say of each that there are infinite) (Galileo, 1954; Parker, 2019, p. 376). Galileo's prioritisation of BP & PWP over relations of size of infinite sets underpins my argument in this essay, that we ought not accept the Cantorian view that PWP must be disregarded in order to improve the functionality of set theory.

Cantor, in pursuit of a useful and enlightening theory of distinct 'absolute transfinite numbers', equates 'equal in cardinal number' with bijection: $(There\ is\ a\ bijection\ from\ A\ to\ B) \Leftrightarrow (|A| = |B|)$ (Cantor, 1932, p. 413).
As demonstrated by Galileo's paradox, we cannot simultaneously maintain PWP and BP along with their application to infinite sets. Consequently, Cantor rejected PWP (Cantor, 1932, cited in Parker, 2019, pp. 375-376).[4] I will now analyse the consequences of Cantor's conception of 'equal in cardinal number'. Firstly, and most notably, this response leads to some counterintuitive conclusions that stem from the rejection of PWP. For instance, Cantor's conclusions imply that there are as many natural numbers as there are integers, despite it seeming, at face value, that there are twice as many integers (or one more, {0}, than twice as many). A more exaggerated version of this counterintuitive result can be seen if we construct a plane, and on that plane, we construct a line segment (illustrated below):



[4] Cantor, of course, allows that PWP is still true for *finite* sets and numbers, but not for *all* sets and numbers.

In this diagram, Cantor's conception of 'equal in cardinal number' compels us to conclude that there are equally many points in the infinite plane as there are in the line segment $\alpha\beta$ (namely, $2^{\aleph_0}$). For brevity, I will not explore what specifically makes something a 'good' or 'sufficient' basis for a conclusion in mathematics; however, counterintuitive results such as this, whilst not sufficient for the refutation of a theory, certainly warrant further critical analysis of the theory. Similarly, whilst elevated blood sugar alone is insufficient for a diabetes diagnosis, it serves as a reason to conduct further tests—just as the counterintuitive results of Cantor's conclusions indicate the requirement for further scrutiny.

Alongside the aforementioned counterintuitive results, Cantor's conception of cardinal equality also appears to lead to some 'inconsistencies'.[5] Some basic operations applicable to sets and numbers allow us to draw general conclusions, such as the following: $A \subseteq B, |B/A| = |B| - |A|; x - x = 0; y \leq (y \times y)$, for sets A, B, integer x and natural number y. Such laws are exceptionally useful in mathematics as well as intuitive. Therefore, if we are to encounter problems applying these laws to 'transfinite numbers', it poses a significant problem for Cantor. This is especially troubling because the ability to work with 'transfinite numbers' is a large factor in considering Cantor's conclusions to be especially useful (Parker, 2008, p. 20; Cantor, 1932).[6] Indeed, when we try to apply these laws to 'transfinite cardinal numbers', we encounter numerous problems, of which I will focus on one example:

Without Cantor's conclusions regarding cardinal equality (and its transfinite consequences), sets and numbers behave in a manner consistent with intuitive expectations:

$$A = \{1,2,3,4\}$$
$$B = \{1,3\}$$
$$C = \{7,8,9\}$$
$$N = \{1,2,3,4,\dots\}$$
$$Z = \{\dots -3,-2,-1,0,1,2,3,\dots\}$$
$$T = N/A = \{5,6,7,8,\dots\}$$

It is clear that:

$$|A \cup C| = |A| + |C| = 4 + 3 = 7$$
$$|A/B| = |A| - |B| = 4 - 2 = 2$$

[5] Here I do not mean that Cantor's conclusions lead to *strict* inconsistencies, but rather that they lead to some peculiar, inconsistent behaviours of some basic operators throughout mathematics and set theory.

[6] Of course, Cantor does address this, but as I will later argue in this essay, Cantor's avoidance of this problem is unsatisfactory.

Etc.

But if we try to treat the infinite sets as we would the finite sets we reach strange, differing results:

$$|N/N| = |N| - |N| = \aleph_0 - \aleph_0 = 0$$
$$|Z/N| = |Z| - |N| = \aleph_0 - \aleph_0 = \aleph_0$$
$$|N/T| = |N| - |T| = \aleph_0 - \aleph_0 = 4$$

This is, of course, avoided by Cantor, by arguing that such operators and laws apply differently to 'transfinite numbers' than they do to finite numbers (Cantor, 1932, pp. 285, 293-294). This introduces my third point regarding Cantor's conception of cardinal equality, namely that Cantor's conclusions require a split concept of number, i.e. a concept of 'number' in general, and 'finite' and 'transfinite' sub-concepts of 'number'. We derive our understanding of 'number' from finite numbers.[7] I do not believe that the concept Cantor puts forward ('transfinite number') should be considered a sub-concept of number, as it has almost nothing in common with 'number' as one would conventionally understand it prior to the introduction of 'transfinite numbers'. As I have shown, operators, rules, and theorems we previously took to be regarding and applicable to (finite) 'numbers' do not apply in the same way to 'transfinite numbers'. This invites examination of what the concept of 'transfinite number' has in common with our basic concept of number that justifies its classification as such. Of course, the Cantorian's reply to this is obvious: 'transfinite numbers enumerate sets', a 'transfinite number' is answering the question 'how many?'; this is a fundamental quality of (cardinal) 'number'. However, the claim that 'transfinite numbers' do, in fact, enumerate sets seems unwarranted. Indeed, in labelling the 'number of natural numbers' as $\aleph_0$ (defining $\aleph_0$ in this way), Cantor is presupposing that there can be such a number (Cantor, 1932, p. 283). Thus, I contend that Cantor is mistaken in his description of 'transfinite number' as a type of number, and that we should not regard 'transfinite numbers' as such as they have very little in common with our original (pre-transfinite) concept of number. 'Transfinite numbers' only enumerate sets insofar as Cantor declares $\aleph_0$ to be the number of members in the natural numbers. This is not a substantial reason to consider $\aleph_0$ to be a number. I propose that, rather than considering $\aleph_0$ to be a number, we should consider $\aleph_0$ to be a property of some infinite sets. For example, we may slightly change our prose to say 'Some set is '$\aleph_0$-like' if it can be put in bijection with the natural numbers (N)' (rather than saying that the set 'has $\aleph_0$ many members').

[7] By this I mean that we initially get a concept of number from a young age through our introduction to the natural numbers. This concept, 'number', can then be expanded to include other types of numbers.

In the face of my critiques of Cantor's conclusions, there is a common defence of Cantor's conception of 'equal in cardinal number' and its implications for 'transfinite numbers': Cantor's view is uniquely useful and enlightening, enabling us to work with and draw conclusions in mathematics—particularly set theory—and to achieve a greater understanding, both of which we could not otherwise reach.[8] This defence initially seems persuasive; however, I will argue that this defence fails to recognise that the functionality of Cantor's set theory can be largely (if not entirely) preserved, whilst avoiding the aforementioned issues with Cantor's theory. As I understand it, the usefulness of Cantor's theory lies in its applications in set theory, particularly in furthering what we can do in set theory by developing laws, theorems and rules regarding sets of distinct 'transfinite cardinalities'.[9] However, it seems to me that we can maintain the usefulness of Cantor's theory without subscribing to the entirety of Cantor's view. Thus, we can keep each theorem or property regarding 'sets of cardinality $\aleph_0$', but, instead of saying they apply to 'sets of cardinality $\aleph_0$', we can simply say that these theorems and properties apply to 'Any set A such that there is a bijection between N and A'. Thus, the notion 'has cardinality $\aleph_0$', or 'has the same cardinal number as the natural numbers' can be removed, without removing the theorems and tools derived from this theory.[10] For example, consider the following theorem: The union of any two sets of cardinality $\aleph_0$ itself has the cardinality $\aleph_0$ (or '$(|A| = |B| = \aleph_0) \rightarrow (|A \cup B| = \aleph_0)$'). This theorem can simply be reformulated as follows: The union of any two sets that can be put in bijection with the set of all natural numbers (N) can itself be put into bijection with the set of all natural numbers. This can be rewritten as follows: $(F: A \rightarrow N, G: B \rightarrow N. Where\ F, G, are\ bijective) \rightarrow (H: (A \cup B) \rightarrow N, for\ some\ bijective\ H)$. Here we preserve the theorem but distance ourselves from Cantor's notion of distinct *greater*, *lesser*, or *equal* 'transfinite numbers'. Furthermore, in Wittgenstein's *Remarks on The Foundations of Mathematics* (*Remarks*, henceforth), Wittgenstein

[8] The *usefulness* of Cantor's conclusions is both more common and more substantial a defence of his conclusions than their supposed enlightening nature; Thus, I will primarily address the claim of usefulness in this essay. Additionally, rather than a particular defence of Cantor's conclusions, this is rather a representation or construction of a sentiment shared by multiple scholars writing on set theory. (Parker, 2019, p. 377; Easwaran, 2024, §3.3)

[9] I do not mean to claim that Cantor's theory of equality of cardinal numbers is exclusively useful for this purpose. However, this is certainly a significant area in which Cantor's conclusions prove useful, and perhaps the primary justification for adopting Cantor's conclusions. Specific applications of Cantor's conclusions will be discussed later in the essay.

[10] This is generalisable to Cantor's other 'transfinite numbers' by speaking of them as 'in bijection with' some other infinite sets, e.g. the real numbers.

expands on this claim, arguing that there is no employment of Cantor's statements as he writes them, but rather these employments are to be invented (Wittgenstein, 1956, p. 59e).[11] Putnam, in response to Wittgenstein (and indeed in a response broadly relevant to my own claims), points out the apparent usefulness of Cantor's formulation of his conclusions, specifically highlighting Major Fixed Point Theorems of recursion (Putnam, 2007, p. 243).[12] This example indeed makes use of Cantor's conclusions and language, but it does not *require* this Cantorian viewpoint, as the instances of 'transfinite numbers' can be treated as I have previously outlined. Cantor's claims are indeed useful, but so too are they troubling, and if we can maintain their usefulness whilst avoiding the relevant troubles, we ought to. Furthermore, as pointed out by Wheeler, Wittgenstein neither claimed any necessary 'uselessness' of Cantor's language, but rather claimed that its employment is "still to be invented" (Wittgenstein, 1956, p. 59e; Wheeler, 2022, p. 327).

I will briefly comment on the notion that Cantor's conclusions are particularly 'enlightening'. This notion is formally refuted to some extent in Wittgenstein's *Remarks*, wherein he demonstrates that Cantor's statements are not statements of fact, but rather determinations, that is, formations of a concept (Wittgenstein, 1956, p. 56e).[13] Wittgenstein contends that, whilst certain considerations that lead us to Cantor's conclusions can be made, we need not make such considerations and indeed it seems decidedly useless to do so.[14]

[11] Wittgenstein claims that "[an] expression is empty so long as there is no employment for it" (Wittgenstein, 1956, p. 59e). That is to say that there was no application of Cantor's concept of 'transfinite number' at the time that this concept was determined. The applications of Cantor's conclusions as discussed in this essay were "still to be invented" at the time of Cantor's writing (Wittgenstein, 1956, p. 59e).

[12] Though an extended treatment of Putnam's response to Wittgenstein eludes the scope of this essay, we may focus here on Putnam's claims that Cantor's conclusions are indeed (seemingly uniquely) useful, and that their usefulness is an important factor in considering their acceptance (Putnam, 2007). Furthermore, in Gödel's writings on Major Fixed Point Theorems of recursion, there are dual uses of the infinite throughout; these being the use of limits (infinitesimals and potentially infinite series), and the use of distinct 'transfinite numbers'. These are acutely different notions and the former is of no concern to this essay. It is apparent that the latter is utilised throughout Gödel's essay. This can serve as an example of the application of my prior comments, i.e. it is apparent that the conclusions reached in Gödel's essay can be reached using the language or 'prose' of properties of infinite sets (such as their being in bijection with the natural numbers) rather than their having particular 'transfinite cardinalities' (Gödel, 1992; Putnam, 2007, p. 243).

[13] I use both 'determination' and 'formation' here in accordance with Wittgenstein's own usage (Wittgenstein, 1956, p. 56e).

[14] Wittgenstein's claim here is dependent on his wider conception of the ontology and epistemology of mathematics. A thorough examination of these foundational conceptions lies beyond the scope of this essay.

Furthermore, the claim that Cantor's conclusions are enlightening is entirely dependent on the acceptance of his conclusions. That is to say, if a conclusion is true, it is of course enlightening to proclaim said conclusion, but if the conclusion is not true the proclamation is no longer regarded as enlightening. As a result, a particular view being enlightening is entirely dependent upon its truth, and thus the argument that we should accept a claim based on its enlightening nature is entirely question begging and void of any real persuasive weight. As such, to reach the claim that Cantor's conclusions are particularly enlightening, one must presuppose these conclusions to be true. Therefore, defending Cantor's conclusions on the basis of their supposed enlightening nature is unpersuasive, and, as such, I reject this defence.

Before concluding, I will briefly outline some differences between my claim in this essay and Wittgenstein's claim in his *Remarks*, as they are significantly similar yet differ on some key aspects. I ultimately argue that Cantor's conclusions ought to be reformulated because they are not uniquely functional; they are not particularly enlightening; they are misleading; they lead to a great deal of change and upheaval of prior mathematical rules, concepts and language use in order to avoid contradictory conclusions. Therefore, on the grounds of theoretical parsimony, we ought to reformulate Cantor's conclusions, and revise the language used in set theory and in mathematics. Alternatively, Wittgenstein ultimately argues that Cantor's conclusions would be unproblematic if they were framed as determinations, that is, formations of concepts, rather than discoveries of facts of nature (Wittgenstein, 1956, p. 59e). Ultimately, my claim is of smaller scope: We have ample reason to avoid the use of Cantor's conclusions and theories as originally formulated as they lead to many problems that can be avoided without losing the benefits of Cantor's system. This differs from Wittgenstein's claims regarding mathematics and mathematical prose in general (Wittgenstein, 1956).

In conclusion, Cantor's account of cardinal equality yields undesirable results in mathematics, particularly when applied to infinite sets. Nevertheless, it is a useful theory in mathematics. I have demonstrated one way to preserve the functionality of Cantor's theory whilst mitigating its problematic implications. These considerations, taken together, warrant dismissing Cantor's conclusions as presently formulated and pursuing a similarly functional, less problematic alternative formulation and understanding of the infinite and cardinal equality.

Bibliography:

Cantor, G. (1932). *Gesammelte Abhandlungen*. Berlin: Springer.

Easwaran, K., Hájek, A., Mancosu, P., and Oppy, G. (2024). 'Infinity.' In E. N. Zalta and U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Summer 2024 Edition). Available at: https://plato.stanford.edu/archives/sum2024/entries/infinity/. (Accessed 25 November 2024).

Galilei, G. (1954) [1638] *Dialogues Concerning Two New Sciences*. (H. Crew and A. de Salvio, Trans.). New York: Dover.

Gödel, K. (1992) *On Formally Undecidable Propositions of Principia Mathematica and Related Systems*. New York: Dover Publications.

Parker, M. W. (2019) 'Gödel's Argument for Cantorian Cardinality.', *Nous* 15: 375–393.

———. (2008). 'Philosophical Method and Galileo's Paradox of Infinity.', London School of Economics, *PhilSci Archive* [online]. Available at https://philsci-archive.pitt.edu/4276/ (Accessed 25 November 2024).

Putnam, H. (2007) 'Wittgenstein and the Real Numbers.', In A. Crary (Ed.), *Wittgenstein and the Moral Life: Essays in Honor of Cora Diamond*. Cambridge: The MIT Press.

Wheeler, S. J. (2022) 'Defending Wittgenstein's Remarks on Cantor from Putnam.', *Philosophical Investigations* 45(3): 320–333.

Wittgenstein, L. (1956) *Remarks on the Foundations of Mathematics*. (G. H. von Wright et al. Ed.) Translated by G. E. M. Anscombe. Oxford: Basil Blackwell.

Are Counterfactuals Enough? Assessing Williamson's Defence of Modal Metaphysics

By Cathal Lacey

Counterfactual modality, as a field of study, pertains to what might have otherwise been. At first glance, modality might strike the reader as a strange subject for philosophy to concern itself with. Afterall, to adopt Lewis' famous example, who cares whether kangaroos topple when their tails are lopped off? However, many philosophers believe that scientific pursuits are modally grounded, and thus that modality ought to play an important role in philosophical discourse. Lacey explores this assumption and gets to the very centre of this debate. He draws on the work of Williamson, an influential contemporary thinker, and giants of the field like Quine to craft this elegant paper.

Introduction

Underlying the metaphysical debate between Modal Realism and Modal Anti-Realism is a concern over whether or not we are justified in conducting investigations into metaphysical modality and its constituent elements, metaphysical necessity, contingency and possibility. A notable perspective on these matters is Metaphysical Naturalism, which proposes that Metaphysics ought to be closely aligned with science and should not deviate from science's best theories (Emery, 2023). Thus, to justify or reject metaphysical investigations into modality, Metaphysical Naturalists attempt to evaluate how strongly metaphysical modality relates to science and its practices. This essay will evaluate the metaphysical naturalist Timothy Williamson's methodological argument that the use of counterfactual reasoning in science provides sufficient justification for the philosophical investigation of the metaphysical modality and necessity. To achieve this aim, it will first briefly outline metaphysical naturalism and the requirements that it places on metaphysical theorising. It will next give an extended account of Williamson's argument in defence of modal metaphysics, before responding that it is not convincing as it violates the standards of metaphysical naturalism. Finally, it provides a more general critique of metaphysical modality, positing that without clear reductionist movement towards knowledge of the mechanisms and joints of reality, modal metaphysics will have little claim to be creating actual knowledge about the true nature of modal reality. Consequently, Williamson's methodological assertion that counterfactuals offer us insight into metaphysical necessity is not convincing.

Metaphysical Naturalism

Among other reasons, given the great success of natural science in mapping the world and reality, William Van Orman Quine held that the philosophical pursuit of worldly knowledge ought to be subject to the "fundamental conceptual scheme of science" and its standards of justification and evidence wherever possible (1960, p. 276 cited in Hylton & Kemp, 2023). He tells us that we must recognise that "it is within science itself, and not some prior philosophy that reality is to be identified and described" (1981, p. 21). Rather than applying a purely a priori metaphysics, Quine proposed that we adopt a naturalised metaphysics which describes "the true and ultimate structure of reality" (1960, 221) by "distilling the essence of scientific discourse [and] profitably purifying the language of science beyond what might reasonably be urged upon the practicing scientist." (1957, p. 235). Williamson, who is also a metaphysical naturalist, tells us that a metaphysical theory of modality will appear "a scam" if it can "be surgically removed from our conceptual scheme without collateral damage" to our scientific and ordinary reasoning (2007, p.136). Based on the above account, we will propose the following criteria as requirements for the success of Williamson's argument: 1) it must align itself with the standards of justification of science, 2) removing it from our conceptual scheme must cause collateral damage in our scientific reasoning, and 3) it must align itself with the fundamental conceptual scheme of science. While Williamson and others might discuss and posit a large range of metaphysical naturalist criteria, these three provide us with an intuitive and reasonable naturalistic baseline by which we can evaluate his claims.

Williamson's Defence of Modal Metaphysics

Williamson's defence of the methodology of metaphysical modality has two parts (Vetter, 2016). First, he identifies counterfactual reasoning as an entry point into modal thought by pointing to its applications in everyday cognition and scientific practice. Second, he moves beyond this initial entry point into claims about wider knowledge of metaphysical modality. This second step begins with the proposition that the use of counterfactuals in science presupposes that science holds nomic modal commitments. Next, he argues that these nomic commitments presuppose metaphysical modal commitments. Thus, Williamson argues that the use of counterfactuals in science presupposes metaphysical modal commitments, and therefore — to take into account these commitments — second-order quantified modal logic should be the logic of science. Based on this line of argumentation, he claims that we are justified in conducting investigations in metaphysical modality through second-order quantified modal logic. We will now move through his reasoning in more detail.

Williamson defends the metaphysics of modality by arguing that modal talk is not "some sort of illicit projection or unacknowledged fiction" that is "the exclusive preserve of philosophers" (2007, p.135). Instead, he argues that counterfactual reasoning plays important roles in scientific practice, primarily in scientific causal reasoning. To illustrate the relationship between counterfactuals and causal reasoning, Williamson presents the following example where we investigate a potential causal relationship between variables X and Y (Williamson, 2016). We set X=1 at time t and observe Y=0 at t+1. With this information alone, we cannot confirm that X=1 caused Y=0. Now suppose we have a non-modal mathematically formulated scientific theory proposing that Y=1–X. Based on this theory, we can counterfactually expect that if it had not been the case that X=1, then it wouldn't have been the case that Y=0. Therefore, counterfactual reasoning provides us with evidence for causation within the context of this theory. This approach is particularly valuable in scientific contexts where controlled experiments are impractical or longitudinal data is limited. He posits that each successful scientific application of such reasoning abductively supports counterfactual thinking more generally. Thus, Williamson rejects the claim that counterfactuals are merely speculative metaphysical tools; rather, they are integral to the logical and mathematical frameworks scientists use to test, refine, and validate causal theories.

In the second step of his reasoning, Williamson (2016) argues that the use of counterfactuals in mathematics and science presupposes nomic and metaphysical modal commitments and thus provides us an avenue into wider knowledge of metaphysical modality. Williamson first highlights the distinction between Objective Modalities and other modalities (epistemic, deontic, and teleological). Objective modalities describe the structure of reality itself, independent of human knowledge, morality, or purposes. While the opposite is true for epistemic, deontic, and teleological modalities. Two primary objective modalities are relevant here: nomic modality and metaphysical modality. Nomic modality closely relates to science and refers to what is possible or necessary given the laws of nature in our world. Metaphysical modality is broader and concerns all possibilities and necessities, regardless of whether they conform to our universe's laws.

To show that mathematical and scientific theories themselves have metaphysical and nomic modal commitments, Williamson presents us with a scientific theory T, which supports counterfactuals (>), obeys the classical laws of physics, and has the truth-functions ¬ and → (material conditional). The premise that T supports counterfactuals can be formalised as follows:

(P1) Whenever T entails F → G, T supports F > G.

Now assume that T entails F and that P1 holds. From the rules of classic propositional logic, we can derive from F the tautology that ¬F → F. This follows because an implication is true when its consequent is true or its antecedent is false, both of which hold in this case. By the transitivity of entailment, we can then infer that T supports ¬F → F; then from P1 that ¬F > F. In other words, ¬F is impossible or □F (where □ is the operator for metaphysical necessity and ◊ is the operator for metaphysical possibility). If ¬F were possible, the counterfactual ¬F > F would not hold. Thus, for ¬F > F to hold, ¬F must be impossible. Therefore, the theory T supports the metaphysical necessity of F:

(P2) Whenever T entails F, T supports □F.

If T is a theory of pure mathematics, this conclusion does not seem overly problematic (given that mathematical statements are considered true in all possible worlds). However, Williamson acknowledges that P2 seems too strong for the theories of natural science, which do not assert that their propositions hold in all possible worlds (metaphysical necessity); instead, just that they hold under the actual laws of our universe (nomic necessity). Thus, for scientific theories, he weakens P1 to only apply to nomic necessity:

P1* Whenever T entails F → G, T supports that ◊nF → (F > G).

Here, if T entails F → G, we can infer that if F were nomically possible (◊n), then the counterfactual F > G would hold. Similarly to his above move from P1 to P2, Williamson moves from P1* to P2*, concluding that the theory T supports the nomic necessity of F:

P2* Whenever T entails F, T supports □nF.

Based on the above two arguments, Williamson proposes that purely non-modal mathematical theories which use counterfactual reasoning have metaphysical modal commitments, and similarly, scientific theories have nomic modal commitments.

Williamson now builds on his claim that science has nomic modal commitments by arguing that scepticism about metaphysical modality generalises to nomic modality (2016). As they stand or fall together, if we are to coherently hold nomic modal commitments about science, we must also hold metaphysical modal commitments. Williamson raises two sceptical arguments that he claims generalise across objective modalities: epistemic and coarse-grained

criticisms. As we might guess from its title, the epistemic criticism questions the basis of our knowledge of metaphysical modal truths. If a state of affairs is non-actual, how do we know whether it is possible? For example, how can we justify the claim that it is metaphysically possible for water to freeze at temperatures other than 0°C in normal conditions, given that this does not occur in the actual world? This epistemological challenge questions the criteria we can use to determine whether a possible or necessary scenario is genuinely possible or necessary. However, Williamson argues that the same problem can be raised when we justify nomic claims about possibility and necessity when considering scenarios that do not occur in the actual world. For example, consider the theoretical scientific question about whether it is nomically possible for gravity to repel instead of attract. Under the lights of the epistemological criticism, we cannot have knowledge about this question as it contradicts everything we have observed in the actual world. The second generalisable criticism is that of coarse-graining, where requirements for modal claims are too easily satisfied. If metaphysics is overly permissive about truth in all possible worlds (metaphysical necessity), it can be trivially easy to claim that certain propositions are necessary (Fine, 1994). While nomic necessity is narrower than metaphysical necessity, claims of necessary nomic truth can also be considered overly cheap. Consider that supposedly nomic claim that 'a dropped object will fall due to gravity', which might be nomically necessary, but does not reflect the complexity and nuance of how gravity operates under different physical conditions.

Based on the generalisability of these two criticisms, he argues, given that if "metaphysical modality is in trouble, so is nomic modality" and that nomic modality plays an important role in science, we should accept a commitment to metaphysical modality (2016, p. 463). The use of counterfactuals in science presupposes nomic modal commitments, and nomic modal commitments demand metaphysical commitments; therefore, by transitivity, metaphysical commitments follow from the use of counterfactuals in science. Williamson thus argues that to account for the modal commitments of science, we should adopt second-order quantified modal logic (QML) as the logic of science. Leaving out the details for the sake of brevity, second-order QML is essentially first-order QML, which is enhanced with quantification over properties (Divers, 2024). Thus, its adoption will allow us to reason formally in science with and about modal properties. This proposal is reinforced by the fact that second-order QML can be developed in terms of counterfactual logic (Williamson, 2016), which, as argued above, plays an important role in science. In sum, Williamson argues that given that counterfactuals are endogenous to science and that the logic of science ought to be second-

order QML—which is an important logic in metaphysical explorations of modality—we are justified in conducting investigations into metaphysical modality.

Evaluation of Williamson's Argument

In order to evaluate Williamson's argument, we will return to the three highlighted metaphysical naturalist requirements placed on metaphysical theorising. Based on these criteria, the most problematic move in the above reasoning is the claim that nomic modality stands or falls with metaphysical modality. It is unclear why we would be convinced that the two critiques of metaphysical modality, which Williamson argues generalises to other objective modalities, would significantly impact our claims about nomic modality. Claims about nomic necessity and possibility are justified based on empirically testable mathematical formulations of physical laws, not broader metaphysical and philosophical coherence. Consider the laws of thermodynamics, which are well established and clearly define constraints on what is nomically possible and necessary. For example, based on the second law of thermodynamics (which prohibits decreases in entropy in a closed system over time), we can state that a perpetual motion machine is impossible given the laws of nature of our world (Deltete, 2012). It is entirely unreasonable to argue that our ability to coherently make this claim of nomic impossibility hinges on the unresolved philosophical issues Williamson puts forth. Secondly, claims of nomic necessity are validated by the stableness and robustness of physical laws over time. Scientific theories like Newtonian mechanics—which has demonstrated consistent robustness (within specific scales) across centuries—are clearly insulated from metaphysical issues that have only matured in the last sixty years (Kment, 2015). Thus, contrary to the first demand, Williamson's claims fail to align themselves with the standards of justification of science.

Furthermore, even if we were to be convinced that claims of nomic modality were tied to metaphysical modality, it remains unclear why this should affect our understanding of the laws governing hard sciences. Williamson himself cautions that modal claims must demonstrate relevance, avoiding metaphysical labelling akin to a philosopher going around "sticking gold labels on [their] favourite machines, reading 'This machine has been approved by a qualified metaphysician' " without making any effect on how the machine operates (2016, p. 465). Descriptions of nomic modality are inherently pragmatically oriented, focusing on mapping regularities in the behaviour of physical properties of entities (charge, mass, location) given the known physical laws. Scientists do not need to decide whether physical laws are nomically contingent or necessary to use them effectively. For example, metaphysical

debates about the nomic possibility of faster-than-light travel do not impinge upon the predictive regularity described in the formulae which underlie general relativity. In response to this issue, Williamson highlights Dynamic Systems Theory as an example where modal talk is an "endogenous need of the science" (*ibid.*). Here, he provides an extended account of the similarities between the metaphysical possible worlds theory and the scientific Dynamic Systems Theory, which culminates in the following claim:

> *In metaphysics we may reason about modality by quantifying over possible worlds; in [dynamic systems theory] we reason about modality by quantifying over possible states of a physical system. In the former case, the relevant modality is metaphysical; in the latter, it is more like nomic (Williamson, 2016, p. 479).*

Even in these rare cases where modality is relevant in science, it is improbable that a specialist in dynamic systems theory would ever be impacted in his research by philosophical concerns over whether his theories presuppose nomic or metaphysical modal commitments. Contrary to the second demand, removing metaphysical modality from our conceptual scheme is unlikely to cause any collateral damage to scientific reasoning.

Finally, based on the third highlighted demand of metaphysical naturalism, those who seek a metaphysical account of reality grounded in science ought to look not only to science's findings and methods but also to its principles and conceptual scheme. Even if they remain open to refinement and debate, metaphysicians who truly want to align themselves with science should place great weight on the principles of empiricism, reductionism, realism, naturalism, positivism, materialism, and falsificationism. One principle stands out in this context: reductionism. Complex phenomena, like modality, should be understood by breaking them down into simpler parts. Without movement towards clear knowledge of the mechanisms and joints of reality, metaphysics will have little claim to the actual creation of knowledge about how reality objectively works (Yudkowsky, 2018). Metaphysicians have delved in great depth into the concepts of necessity, possibility and contingency, producing a vast number of modal logics with which they claim insight into the mysterious realities of modality. Each new modal logic is another way of talking about the mystery, putting new labels on the mystery, and creating new means of manipulating the mystery without actually moving closer to the gears of the mystery. If real progress is to be made in our knowledge of modality, we should first turn to cognitive science. Instead of asking what the term 'necessity' means and what it refers to in the world, we should be asking what cognitive algorithm creates the inner perception of necessities. One might begin by looking at Gibson's (1979) affective psychological theory of perception, in which people perceive the world in relation to possible

action paths. It is not an unreasonable possibility that Williamson's modal metaphysical talk is the product of erroneous abstraction of a non-ontologically fundamental element of our cognition and perception. We should not take the fact that we perceive and reason about possibilities at face value. It is wrong to assume that these capacities give us direct access to the structure of reality itself. Only once cognitive science can uncover the cognitive mapping algorithms that relate to possibility and necessity will we be able to separate what originates in our minds from what emerges in reality itself. Discussion of a particular form of reasoning—counterfactuals—in metaphysics is not likely to facilitate our reduction of necessity down to its gears, nuts and bolts. Because of this failure to enforce a reductionist methodology, Williamson cannot claim to align itself with the conceptual scheme and principles of science.

Conclusion

In conclusion, Williamson's methodological defence of metaphysical modality in terms of counterfactuals does not meet the posited standards of metaphysical naturalism. Williamson appears to be doing precisely what metaphysical naturalism says he should not do. From his philosophical armchair, he argues that due to unresolved philosophical issues, we and scientists could not coherently make nomic modal claims about our scientific theories if we did not also accept metaphysical modal commitments. The fundamental idea of metaphysical naturalism is that metaphysicians ought to look to science when describing reality, not vice versa. If, however, his argument is not as prescriptively overreaching as I claim, then we must again ask does it meaningfully affect scientific reasoning, as required by the second criterion. Even in rare cases where questions of modality are pertinent to the practice of science, claims of nomic modality should be judged as consistent and meaningful in reference to the robustness of scientific theories; not with the metaphysics of modality grounded on counterfactual reasoning. Consequently, Williamson does not provide a convincing methodological proposal that counterfactuals offer us insight into the metaphysics of modality and necessity.

Bibliography:

Deltete, R. (2012) Planck, Ostwald, and the Second Law of Thermodynamics. *The Journal of the International Society for the History of Philosophy of Science, 2(1), pp.121-146.*

Divers, J. (2024). Seminar 9: Williamson on Counterfactuals and (Metaphysical) Necessity. PIU44041: Metaphysics. 21st Nov.

Emery, N. (2023) *Naturalism Beyond the Limits of Science: How Scientific Methodology Can and Should Shape Philosophical Theorizing*. Oxford University Press.

Fine, K. (1994) 'Essence and Modality: The Second Philosophical Perspectives Lecture,' *Philosophical Perspectives*, 8(1).

Gibson, J.J. (2014). *The Ecological Approach to Visual Perception: Classic Edition* (1st ed.). Psychology Press.

Hylton, P. & Kemp, G. (2023) Willard Van Orman Quine, *The Stanford Encyclopedia of Philosophy* (Fall 2023 Ed.).

Kment, B. (2015) *Modality, Metaphysics and Method,* in Daly, C. (2015) *The Palgrave Handbook of Philosophical Methods*. Palgrave Macmillan.

Quine, W.V.O. (1957) The Scope and Language of Science*, British Journal for the Philosophy of Science*, 8(29), pp. 1–17.

Quine, W.V.O. (1960) *Word and Object*. MIT Press.

Quine, W.V.O. (1981) *Theories and Things*. Harvard University Press.

Vetter, B. (2016) 'Williamsonian modal epistemology, possibility-based,' *Canadian Journal of Philosophy*, 46(4–5), pp. 766–795.

Williamson, T. (2016) 'Modal science,' *Canadian Journal of Philosophy*, 46(4–5), pp. 453–492.

Williamson, T. (2007) *The Philosophy of Philosophy*. John Wiley & Sons.

Yudkowsky, E. (2018) *Map and Territory*. Machine Intelligence Research Institute Press.

"There are holes in that cheese" —Revisiting Lewis and Lewis on Holes

By Gavin Dunphy

I hope you're hungry! This delicious paper is playfully written, but nevertheless takes up one of philosophy's more difficult metaphysical problems. I expect many readers will not have thought much about the existence, or lack thereof, of holes but the following paper does just that. As the reader will discover, holes present something of a unique issue for philosophers. They are inherently difficult to define given that they are marked by a lack, rather than presence, of a particular substance. To address this, Dunphy grapples with influential metaphysical works by Quine and Lewis and Lewis in attempting to find a position that is both intuitive and unobjectionable.

Entrée: What is physicalism?

In this paper, I will argue that it is possible to be a physicalist and make the claim that "there are holes in that piece of cheese", in response to Lewis and Lewis' charming paper 'Holes' (1970). I will argue for an understanding of physicalism that renders it compatible with a grounding position on meta-ontology based on Jonathan Schaffer's view, and then explain and argue for that view against a Quinean position on ontology. I will then show that the existence of holes, insofar as they are said to exist in the claim "there are holes in that piece of cheese" is compatible with physicalism by arguing that they can exist as dependent entities within Schaffer's framework, and owing to their ontological dependence on fundamental physical entities can properly be called 'physical entities'. This is intended to show that an alternate route is possible to that depicted in the Lewis and Lewis paper. I will then deal with the objection that this framework leads to entities that are vaguely existent, arguing that on the framework as it is used here, this is tolerable.

First, it is important to argue for a certain understanding of physicalism. We will take it that the best understanding of physicalism is not that 'only physical entities exist', but that 'whatever exists fundamentally is physical' (whatever physical means). By 'exists fundamentally' we mean whatever exists such that it is not ontologically dependent on some other entity. These two statements of physicalism, it should be noted, are compatible with each other, because if all existent entities are fundamentally dependent on what is physical, it is reasonable to understand them all as 'physical entities'. The clarification is necessary as the very topic of this paper is the existence of things that are not fundamental at all, and thus not

straightforwardly physical. In this paper we aim to show that holes can be thought of as entities dependent for their existence on physical entities, marrying physicalism with a commitment to the existence of holes. If physicalism were only the doctrine that all that can be understood to exist is fundamentally physical entities, then the claim of this essay would be trivially false because holes are not fundamental particles, or fields, or any other likely contender for what 'physical' means. That our working definition of physicalism is what we should mean by physicalism in discussions like these is evident in the pursuit of a physicalist account of consciousness, and other 'higher level', emergent, or ontologically dependent things. A physicalist philosopher of mind does not think that our experience of colour is a fundamental physical entity, they think it is somehow constituted by physical entities, or explicable in terms of them (Stoljar, 2024).

The Main Course: A Hierarchical Grounding Ontological Framework

Before arguments over the cheese, we must fill our bellies with the main—a "neo-Aristotelian framework, built around primitive grounding relations" (Schaffer, 2009, p. 347). This framework, as expounded by Schaffer, is a valiant attempt to revive a hierarchical ontology against a Quinean flat ontology. We will briefly consider this other option on the menu. Quine would be uncomfortable with strange things like holes and would try to find a way to exclude them from his list of things that exist, usually by paraphrasing statements containing undesirable entities so that these entities instead become predicates (Quine, 1948). In the case of holes, he will likely commit only to matter that constitutes the hole, not to the hole itself as an entity, because he wants to commit only to what our best theories require quantifying over. Instead of committing to the existence of holes, Quine could simply have a predicate meaning "has a hole in it", or "is perforated" holding that this is really what we mean when we refer to something as a hole (Quine, 1948; Lewis and Lewis, 1970, p. 206 f.). This is exactly the strategy that Argle, one of the two fictional characters in the dialogue of Lewis and Lewis (1970), uses at first. The problem with employing this strategy to resolve the difficulty which is the topic of our paper is also demonstrated in Lewis and Lewis (1970), by the other character, Bargle. He asks, "how will you say that there are as many holes in my cheese as crackers on my plate?" (Lewis and Lewis, 1970, p. 207). Argle cannot answer because if there are two holes in the cheese, he is using a predicate like "doubly perforated", whereas if there are two crackers on the plate, he is committing to the existence of two crackers. He thus cannot straightforwardly say that the number involved in the two statements is the same without committing to the existence of two holes in the cheese—this strategy is so impractical as to make pursuit of it thoroughly unconvincing. As Argle goes on to do, we will avoid this

by going for the option that "there *are* holes" (Lewis and Lewis, 1970, p. 207), but we will avoid the pitfalls he runs into by clarifying a viable ontological framework for them first.

Schaffer's ontology consists of (1) fundamental entities, (2) entities which are derivative from fundamental entities, or from other derivative entities, and (3) the link between derivative entities and what they derive from, which Schaffer calls 'grounding'. This relation (3), he says, "passes every test for being a metaphysical primitive worth positing. It is unanalysable. It is useful. And it is clear what we mean" (Schaffer, 2009, p. 376). What he means by it is something like "ontological dependence, priority in nature" (Schaffer, 2009, p. 373). Grounding is the relation between what is fundamental and what is derivative, *derivative entities* are those entities which, were it not for the entities they are grounded in, would not exist, and *fundamental entities* are those which are not grounded in anything. This is the metaphysical framework our argument builds on. The argument, like arguments for all very broad metaphysical positions like this, is that it makes the most sense of the data we want to account for and the problems we want to avoid. The efficacy of the hierarchical grounding ontology will be demonstrated in the reconciliation of physicalism with a desire to truthfully make claims like "there are holes in that piece of cheese". What we are looking to do in pointing to the existence of holes, and not committing to anything other than physical entities, is some kind of preservation of everyday ontology. This is quite a different project from Quine's relentless penchant for parsimony. Quine's motivation is one version of Ockham's Razor, avoiding committing to the existence of as many things as possible (Baker, 2022). Our motivation is different, it avoids *not* committing to things that it really seems we ought to. We are motivated to find a simpler explanation for ontology, rather than just a very parsimonious ontology. This is a goal which Schaffer says requires the revival of traditional metaphysical concerns, about *how* things exist, "a revival of the traditional Aristotelian view, which involves concepts one will not find in Quine or Carnap" (Schaffer, 2009, p. 354).

Pausing for Thought: What is a Hole?

Having swallowed this framework, we take a step outside to clear our heads. We look forward to the cheese, eager to commit to its existence and to that of any holes it might bear, all grounded in physical fundamentals we are proud to continue to hold as fundamental. We start to wonder what a hole actually is. We saw Gruyere on the menu, French style, so we know that if it arrives thinly sliced, it will be perforated, and thus have various holes in it. If there is a whole block of it though, we will see some perforations, along with some indentations that do not go all the way through. These are surely holes too, like a hole one digs in the ground. There will also likely be cavities inside the cheese, are they holes too?

How do we know when something is or is not a hole? (Casati and Varzi, 2004, pp. 2–4; Lewis and Lewis, 1999, p. 186) It seems that what is a hole is dependent to some extent on what is considered to be a hole. We could even come up with different definitions of holes, one would be a complete perforation in some entity, another would include holes that are not perforations of something, others would include various variations in various combinations, some rigorously topological, others not. This is all possible, but it seems reasonable to suppose that the reason it is so difficult to settle on a definition (and to decide if something does or does not meet the definition, notably in the 2000 US presidential election (Casati and Varzi, 2021)) is that holes simply are something abstracted from their background by humans. They can be robustly identified in many important cases but are not rigorously definable and not exhaustively identifiable. That this seems to be the case is a good argument that holes are not specific physical entities, and we must conceive of them in the current framework as entities picked out by humans (who have a physical basis), from the physical world. Allowing for this vagueness in the identification of holes and avoiding the need to answer the question of what exact physical basis they have, avoids the problems that Argle gets into in defining exactly where a given hole is or is not (Lewis and Lewis, 1970, pp. 207–11). We allow for this vagueness because our ontology is permissive, holes trivially exist. The question could then become about the way in which they exist, by what route they are grounded (for example: exactly where the vagueness of definition arises and how we should understand this), but that need not concern us in the current paper. Everything that could be involved in a hole's existence: cheese, the mind, air inside the hole, also trivially exist and clearly have a physical basis. The exact dependency structure of holes is not required for this argument, just that this framework is permissible, and that it entails that holes are physical entities. We also avoid committing to one of a range of peculiar positions on what holes are (Casati and Varzi, 2023), such as tricky "parasitic" relations between holes and hole-linings (Casati and Varzi, 2021, p. 10).

The Cheese Course: An Objection

Returning inside to see cheese on the table, pointing at the Gruyere we say, "there are holes in that piece of cheese", to a companion. She quite reasonably objects that this is an absurd statement given that she knows that we would struggle to define exactly what a hole is and where in the cheese the holes begin or end. Our companion holds that the argument of the last paragraph leads to absurdity, that vagueness is simply impermissible in ontology. Vagueness in language, and maybe in epistemology, is tolerable, but this is not one of those cases. There are no things that ambiguously exist or exist ambiguously.

This objection does have weight against the current framework in a way it does not in Argle and Bargle's argument, where Argle can simply say (as he does in the latter half of the paper) that there are an indefinite amount of different 'hole-linings' incorporating different amounts of the cheese, each one existing straightforwardly, and all of the vagueness of our conception of holes stemming from the ambiguity of our language as to which exact set of straightforwardly existing hole-linings we are referring to (Lewis and Lewis, 1970, pp. 207–11). The present framework will likely admit Argle's many different hole-linings since he has gone to the trouble of describing them, and they do not seem to violate any principles. However, these hole-linings are not what is used to explain the existence of the holes at hand. Our framework is permissive about the existence of most given holes we want to discuss, as 'holes' rather than strange collections of many different potential physical locations. We are trying to explain how we are to understand the way in which holes exist, and how this leads to strange features about them such as vagueness. Our framework must thus commit to the existence of vague holes. The objection is reasonable because in some way the grounding framework contains vague existent entities, and because ontologists are quite reasonably opposed to vague existence.

This objection can be rebuffed through a clarification of how vague entities are to be grounded. We have in our ontology what is fundamental, what is derivative, and the grounding relation between them. It is quite clear that what is fundamental cannot exist vaguely, at least if we are to accept the assumption of the objection, which in this paper we will. It is clear also that the grounding relation cannot be vague because then it would not be a direct grounding relation, there would be slack in the system allowing for input, causal or otherwise, from things not that are not prior in the ontological hierarchy, something we want to avoid in preserving physicalism and the current framework. However, it is not clear that derivative entities must not be vague. Let us say there is a hole in the cheese, and we are asked "is the hole itself the air within this perforation, or does it include cheese as well? If it includes cheese, how much exactly does it include? If there is a hole in the cheese with three openings, is this two holes or one? If I start to melt this cheese, at what exact point will there cease to be a hole?". We declare that there are no exact answers to these questions, though we can answer some of them with more certainty than others. On the first few questions it is clear that we are unsure of *how* the hole exists (i.e. *where* exactly), and on the latter questions it is clear we are unsure what exactly constitutes its existing. In this scenario, though, this does not violate any ontological principles. We can give a rough sketch of how the hole is grounded, it is grounded in cheese, which is grounded physically, and it is grounded in a shared human understanding of what holes are, which is grounded physically. We cannot

know exactly by what route the hole is grounded, but this does not entail that the grounding relations are vague. They are not vague, even those between the hole and whatever is immediately prior to it, if the hole exists (which is a matter concerning the hole), the relations definitely exist. Vagueness in the existence of certain derivative entities like holes seems to arise from things like vagueness of definition, and topological uncertainties. That is all to say, whether or not the hole exists, all of the things that ground the hole exist unambiguously, the vagueness of the hole's existence does not contradict an ontological principle. Whether or not one agrees with this, however, is very much dependent on one's opinion of the whole hierarchical framework.

Digestif

In this paper I have argued that one can be a physicalist, holding that everything which exists is fundamentally grounded in something physical, and make the claim that "there are holes in that piece of cheese". I clarified what it means to be a physicalist and detailed an ontological framework compatible with both physicalism and the existence of holes. I then showed how holes are to fit into this framework, noted that they exist ambiguously within it, showed that this leads to a serious objection to the view, and then countered that view. We conclude that there *are* holes in cheese and elsewhere. What exactly that means is in many cases unclear, but it is more reasonable to support the view that they are grounded somehow, by some route, in the physical, than it is to paraphrase them away. This paper has to some extent provided relief from the incommensurability of the positions that Lewis and Lewis (1970) conclude with, making the space for something of a middle ground between Argle and Bargle (Lewis and Lewis, 1970, pp. 211–12).

I'LL TELL YOU, MY GOOD SIR, WHAT PEACE IS:
THE HOLE WHEN ALL THE CHEESE IS GONE.
(Brecht, 1941, p. 51)

Bibliography:

Baker, A. (2022) 'Simplicity', Section 2, in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*. Summer 2022. Metaphysics Research Lab, Stanford University.

Brecht, B. (1961) *Mother Courage and Her Children*. Translated by E. Bentley. New York: Samuel French.

Casati, R. and Varzi, A.C. (2004) 'Counting the Holes', *Australasian Journal of Philosophy*, 82(1), pp. 23–27.

Casati, R. and Varzi, A.C. (2021) 'Ballot Ontology', in Casati, R. and Varzi, A. C., *Non-Being*. Oxford University Press, pp. 139–164.

Casati, R. and Varzi, A.C. (2023) 'Holes', in E.N. Zalta and U. Nodelman (eds) *The Stanford Encyclopedia of Philosophy*. Summer 2023. Metaphysics Research Lab, Stanford University.

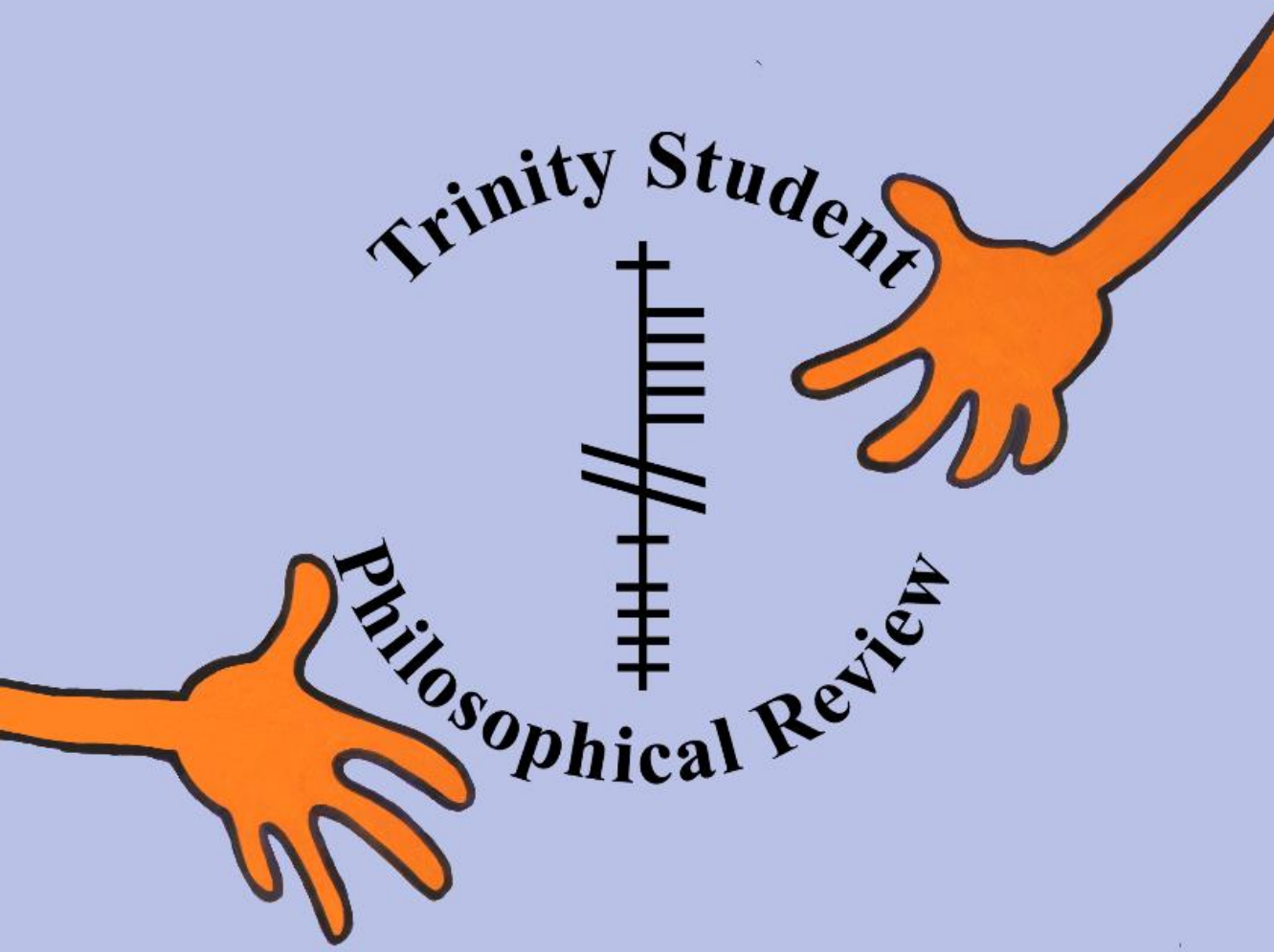
Lewis, D. and Lewis, S. (1970) 'Holes', *Australasian Journal of Philosophy*, 48(2), pp. 206–212.

Lewis, D. and Lewis, S. (1999) 'Casati and Varzi on holes', in *Papers in Metaphysics and Epistemology*. 1st edn. Cambridge University Press.

Quine, W.V. (1948) 'On What There Is', *The Review of Metaphysics*, 2(5), pp. 21–38.

Schaffer, J. (2009) 'On What Grounds What', in D. Manley, R. Wasserman, and D.J. Chalmers (eds) *Metametaphysics: new essays on the foundations of ontology*. Oxford: Clarendon press.

Stoljar, D. (2024) 'Physicalism', Abstract, in E.N. Zalta and U. Nodelman (eds) *The Stanford Encyclopedia of Philosophy*. Spring 2024. Metaphysics Research Lab, Stanford University.


Trinity Student
Philosophical Review