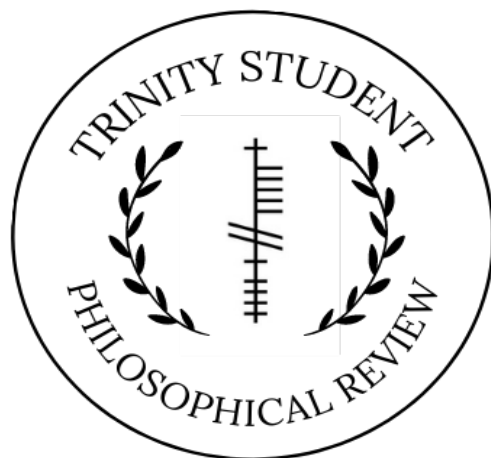# TRINITY STUDENT

PHILOSOPHICAL REVIEW

VOLUME I
2023

Trinity Student Philosophical Review

Patron

Professor John Divers

Department of Philosophy, Trinity College Dublin

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

Tethras Technology Ltd.

Trinity Publications

TCD Trust

http://www.tcd.ie/alumni/services/affinity-credit-card

For Raymond Comyn, Philosopher, Historian, Gentleman.

STUDENT PHILOSOPHICAL REVIEW
COMMITTEE - VOLUME I

Winner of the George Berkeley Award for Best Undergraduate Essay

A Critique of Priest's Master Argument for Dialetheism

Darragh Senchyna

Welcome from the Editor

It is with great pleasure that I welcome you to the inaugural edition of the Trinity Student Philosophical Review. This Review was established to provide undergraduate students with a platform to showcase their work. We hope that the names you read within the pages of this Review go on to contribute to philosophy throughout their lives and based upon the standard of essays we received I think that this is an likely scenario. We have selected ten incredible essays and I hope that there is something for everyone contained in this edition of the Trinity Student Philosophical Review.

The submissions this year were of the highest quality and it was incredibly challenging to select which ones would go to publication. I was truly delighted with both the depth and breadth of the essays we received. The topics covered ranged from the ethical implications of deploying warrior robots into combat scenarios all the way to a deep dive into the logic of sets. Unfortunately, we are unable to publish every submission. The debate around which essays to include was intense and we feel we have included only the strongest essays across a wide range of topics. I thoroughly enjoyed reading each and every submission and without the work of our contributors, we would not be able to create this Review.

I would like to take this opportunity to talk about the essays which we have included in this edition of the Trinity Student Philosophical Review. The Review opens with the brilliantly titled *Error 404 - Morality Not Found?* by Molly Therese Carroll. This essay critically evaluates the error theory and finds that it is unsuccessful in proving that ordinary moral discourse entails a commitment to irreducibly normative objective moral facts and, thus, it ought to be rejected. The following essay: *Robert Sparrow and Artificially Intelligent Warrior Robots* by Taylor Crates, is an interesting dive into the ethical implications of the use of A.I.-powered robots in warfare. This is an especially pertinent essay as the debate around how morality can be applied to artificial intelligence grows more and more relevant every day. The third essay in the Review: *On the Role of Absences in Causal Explanation* by Rían Coady, discusses the role of absences in philosophy. Coady argues that while absences are not causes in themselves they can feature in causal explanations.

The next essay is by Darragh Senchyna and is the winner of our George Berkeley prize for best essay in the Review. It is entitled *A Critique of Priest's Master Argument for Dialetheism* and it truly encapsulates what philosophy in Trinity is all about. The masterful use of logic throughout the essay to come up with a truly unique critique of dialetheism showcase the brilliance of up-and-coming philosophers here at Trinity. It is a well-deserved winner of the inaugural award. Following this is Frank Wolfe's *The 'Darwinian Dilemma' Revisited: Clarifying the Site of Disagreement between Street and Copp*. A fascinating and unique approach to a traditional problem in metaethics. The sixth essay in this Review is *Objects as Sensory Experience: Berkeley's Contribution to the Discussion on Virtual Reality* by Tadhg Cowhig. This essay takes the philosophy of Berkeley and applies it to the modern phenomenon of virtual reality. This is another essay which may become all the more relevant in the future and in the present makes for an intriguing read.

The subsequent essay *The Possibility of Comprehensible Concepts With Empty Sets: Comparing the Standard Contemporary View with that of Strawson's View* by Jack Palmer is an exceptional example of taking the problem of existential import and using logical reasoning to formulate a convincing solution. Following this is Carla Barry's *A Way into the World: Heidegger and the Ready-to-hand*. This essay is a look at the ontology of Heidegger through his phenomenological perspective on time, being, and action. Iain MacLeod's essay *Are There Such Things as Moral Facts?* follows this and returns to metaethics. MacLeod tackles one of the oldest questions in moral philosophy, utilising the writings of Alexander Miller, G.E. Moore, John McDowell, and Jerry Fodor to do so. The final essay in our review this year is *An Exploration of the Phenomenological Conceptions of Space and Place Through the Lens of an Advancing Modernity: From Virginia Woolf to the Present* by Isabelle Logan. This essay is a phenomenological exploration of space and place through the writings of Virginia Woolf. This essay makes a captivating read and is a brilliant conclusion to a very strong set of essays.
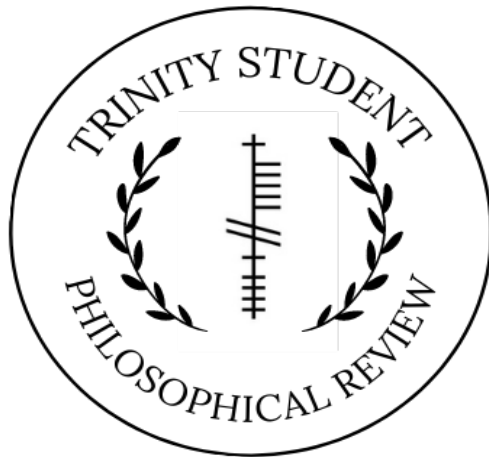
Ódran Farrell
Editor in Chief
Trinity Student Philosophical Review, Volume I

Table of Contents

The Trinity Student Philosophical Review

Error 404 – Morality Not Found?

By Molly Therese Carroll

Preliminary Matters

In this paper, I shall argue that the standard error theory does not succeed in proving that ordinary moral discourse entails a commitment to objective moral facts that are irreducibly normative, and so we should reject the error theory on these grounds. The standard error theory with both its conceptual and ontological claim will be outlined. Finlay's objection to this theory, that moral discourse could have a relational interpretation, will then be considered. I will then examine Joyce's two responses to this objection. Firstly, there could be other presuppositions besides irreducible normativity that demonstrate the systematic falsity of moral discourse. Secondly, moral judgements do, in fact, appeal to objective, and irreducibly normative moral facts. On the first account, I will agree that Joyce's argument has some merit, and consequently will strengthen Finlay's argument to overcome this objection. Then, I will mount a critique of Joyce's argument, and demonstrate how Joyce's second objection fails in disproving this strengthened form of Finlay's argument. Ultimately, I shall conclude that the error theory fails in proving that moral discourse necessitates an appeal to objective moral values that are irreducibly normative, and thus, we should not believe the error theory.

The Error Theory

The standard error theory posits two claims. Firstly, it submits a conceptual claim that in moral discourse, individuals state moral judgements that express beliefs (Miller, 2011, p. 112). This is a cognitivist view because it understands moral judgements as expressing beliefs that are truth-apt; this means that these beliefs can be true or false (ibid. p. 3). They also believe that ordinary moral judgements assume that there are objective moral values (Mackie, 2011, p. 35). For instance, if one were to make the claim that, "you ought not to murder", people would believe that this is objectively true, and ought to be obeyed irrespective of the agent in question's desires or interests (Foot, 1972, p. 307-308). If these beliefs are true, then this would require the existence of objective and prescriptive moral facts to make them true (Miller, 2011, p. 112). Secondly, the standard error theory then submits an ontological claim, that these objective and prescriptive moral facts do not exist (ibid., p. 112). Consequently,

moral judgements cannot obtain the necessary truth conditions needed to make them true, and so the entirety of moral discourse is rendered systematically false (ibid. p. 118).

The Argument from Queerness

Mackie claims that there are fundamental issues with the existence of objective moral facts, and so we should reject their existence. The metaphysical issue he has with them is that they would have to provide an irreducibly normative reason for acting (Miller, 2011, p. 117). This reason would supersede any other reason that could be present and cannot depend on the contingent properties of an individual's desires (Lenman, 2013, p. 401). This would make these facts unlike "anything else in the universe" and therefore queer (Mackie, 1977, p. 49). This property of queerness can be further elaborated when one examines its association with "moral bindingness" (Joyce, 2003, p. 30-31). In moral discourse, we believe agents are bound to act in accordance with moral judgements, even if they have not made any moral judgements themselves (Garner, 1990). For instance, if one were to say, "you ought not to steal", it would not be considered a defence available to a thief that they had never made this moral judgement themselves. Morality is therefore inescapable, which is what makes this relationship between agents and moral facts queer (Joyce, 2011, p. 31).

The argument from queerness and how it relates to both the conceptual and ontological claim can be further understood in the argument below:

P1.   In moral discourse, individuals express moral judgements that entail irreducible normativity.

P2.   If moral judgements entail irreducible normativity, then to meet the necessary truth conditions to render them true, they must commit to objective moral facts.

P3.   Objective moral facts are "queer" because they are morally inescapable.

P4.   If something is "queer", we should reject it.

C1.   Thus, we should reject moral facts.

C2.   Therefore, moral judgements cannot obtain the necessary truth conditions to render them true.

C3.   Moral discourse is systematically false.

In this paper, I will reject premise one. I will argue that irrespective of moral facts being queer, "irreducible normativity" is not an essential feature of moral discourse.

Finlay's Rejection of "Absolute Authority"

Finlay objects to the presupposition of the error theory that moral discourse entails a commitment to irreducible normativity. He offers a relativistic interpretation of moral discourse without this feature (Finlay, 2008, p. 351-352). He understands relativism as the claim that all values are relative to a standard or goal (Finlay, 2011, p. 542). To defend this argument, he examines the evidence Joyce offers to demonstrate the presence of irreducible normativity in moral discourse. Joyce submits that this feature is universally accepted among users of moral concepts. He argues that people accept absolute interpretations of their moral judgements and reject relative interpretations (Finlay, 2008, p. 352). For instance, if a person claims, "you ought not to murder," they would reject another person's interpretation of this being relative to their standards or goals. They would instead claim that this statement should be held in the absolute. This argument, however, does not seem convincing. As Finlay points out, many theories that have been conclusively disproved were once universally accepted. For instance, water was once universally believed to be an irreducible element. However, it is now acknowledged to be a compound of hydrogen and oxygen. Even if this is not true, however, it is not abundantly clear that morality being objective and having an irreducibly normative reason to abide by it, is a universally accepted view. Many people, Finlay included, will favour a more subjective view of morality (ibid., p. 353).

Furthermore, Joyce argues that when we judge an act to be morally good or morally bad, this intuition of irreducible normativity is present because we do not consider if the agent in question's standards were compromised (Joyce, 2003, p. 98). It would be foolish to consider the personal moral standards of Ted Bundy, for instance, when appraising his actions. However, I believe this misunderstands Finlay's relativistic interpretation of moral value. It is not that moral appraisals must be relative to the standards of the person who is being appraised, but rather they can be relative to the standards of the person making the appraisal (Finlay, 2008, p. 353-354). The appraiser's condemnation of an individual's immoral behaviour does not need to consider whether the wrongdoer has committed to their own interpretation of wrongness. We can then all agree that what Ted Bundy did was wrong,

3

according to our own interpretation of wrongness, without needing to worry if this wrongness had any rational authority over him.

Nonetheless, Joyce claims that moral judgements do presuppose that we have an irreducibly normative reason to abide by them. He argues that when we address people who do not subscribe to any moral code, we use objective moral judgements (2003, p. 41-42). This could pose a problem for the relativistic account of moral value, because if the people being addressed do not share the interests, or desires of the speaker, then they do not seem to be relevant. Yet, people will still address these non-subscribers, which seems to suggest that they do not intend the moral judgements stated to be relative (Finlay, 2008, p. 356). Finlay demonstrates this argument is flawed by examining what people pragmatically wish to accomplish by their speech acts in moral discourse. He argues that people express objective moral judgements to serve as a rhetorical aid. Leaving out explicit relativisation makes it appear that the relevant standards are unequivocally subscribed to by the people we address. This conveys the demand for others to subscribe to our own standards, even if they are implicit. The appearance of an objective property in our moral judgements is consequently merely a rhetorical device used to persuade one's audience (ibid. p. 357).

Thus, Finlay's reductionist account of moral values, he believes, demonstrates that having an irreducibly normative reason to abide by moral judgements is not a fundamental feature of moral discourse.

Joyce's Response

Joyce offers a strong defence against Finlay's objections to prove that irreducible normativity is an essential feature of moral discourse. He forwards two main objections. Firstly, he argues that although Finlay is correct in arguing that popular error theories commonly presuppose that moral judgements have an objective value, he is incorrect in assuming that every error theory requires this presupposition. Indeed, there could be many other presuppositions the error theorists could make to prove that moral discourse is systematically false, without appealing to objective value. For instance, moral discourse could be false due to a misconception of human autonomy, or human character traits (Joyce, 2010, p. 520). The only defence Finlay has to offer to this charge is that objective value seems to be the "most familiar and promising candidate", which does not seem convincing as many other theories could be described as both familiar and promising (2008, p. 351).

I would strengthen Finlay's argument here by arguing that even if his objections only apply to these specific accounts of the error theory, this is sufficient as these seem to be the predominant theories and is ultimately the presupposition Joyce makes in his understanding of the error theory as well. However, I do believe that it does seem that many of these other theories do implicitly presuppose that moral claims are irreducibly normative, and so Finlay's objections still stand (ibid). For instance, Anscombe submits that moral concepts used to make sense when they were historically conceived of within a theistic framework. However, now that this framework has been largely removed, these concepts no longer make sense and merely evoke strong intuitions (Anscombe, 1958). Yet, this also seems to presuppose that moral claims are irreducibly normative. These intuitions that originated from theistic frameworks seem to be intuitions of objective value. If we look at most of the dominant religions worldwide, they seem to preach their moral code as objective, and binding to all people. Religious tenets, such as the ten commandments, are not principles meant to apply to only a select few. Consequently, it does seem that different accounts of the systematic falsity of moral discourse can be reduced to the same presupposition that Joyce and Mackie make.

Joyce then forwards another objection which seems to get to the core of Finlay's reductionist account of moral value. He argues that Finlay's relativistic interpretation of morality does not explain how people would respond with disgust or outrage to moral judgements when the speaker explicitly expresses that their view is relative (Joyce, 2011, p. 527). If the judge presiding over the Nuremberg trials were to consistently relativise his appraisal of war criminals by adding the suffix "according to our moral standards", the people present would find this inadequate and would have a strong emotional response. This is because it is not that the people present believe the war criminals' actions to be wrong from their point of view, but that they believe them to be wrong from every point of view. Without this, they feel a strong emotional response because it seems that the sense of wrongness is contingent on their judgements, and would be absent if they were not present (ibid).

Critique

I will put forward three objections to Joyce's defence of "irreducible normativity" as an essential feature of moral discourse.

Firstly, I believe he takes an uncharitable view of Finlay's relativistic interpretation of morality. Joyce understands Finlay's relativism to mean that moral concepts or judgements

are relative to a particular agent (Joyce, 2011, p. 521). Hence, when he discusses the judge at the Nuremberg trials prefacing his condemnation of war criminals with "as per our moral standards", the judge is faced with outrage from the audience present. However, I believe this fundamentally misunderstands Finlay's definition of relativism. He takes relativism to mean relative to any standard or goal (Finlay, 2011, p. 542). This could be but is not required to be an individual's own standards or goals. Consequently, a possible goal could be the maximisation of humanity's happiness. The judge at the Nuremberg trials could then state that what the war criminals did, was wrong, not because we judge it to be so, but because it was objectively wrong in promoting the goal of maximising humanity's happiness (Finlay, 2011, p. 543-544). I do not believe that the audience present would have the same emotional response of outrage to this statement and, thus, Joyce's argument does not seem convincing. Some members of the audience may wish that the judge did state that the war criminal's actions were objectively wrong because this may make them feel more justified in their condemnation, but their desires do not make objectivity an essential feature of moral discourse. To illustrate this with an example, people may wish for water to be a fundamental, non-reductive element but this wishing does not make water any less the compound that it is.

Secondly, even if the previous objection were not true, Joyce's argument still appears implausible. The issue with using an example that requires the presence of war criminals is that it distorts our understanding of what ordinary moral discourse looks like, and how individuals interact with morality. It is not often that an individual is faced with the actions of war criminals and required to condemn them. Usually, the stakes are much lower. Thus, in most situations where moral condemnation is required, a relativistic interpretation of morality does not seem as counterintuitive as Joyce claims it to be. Suppose, for example, a teacher chastises a student for cheating on an exam. She could claim, "cheating is wrong", and the people present would agree. However, she could also claim, "cheating is wrong, according to our school's policy". Relativising her judgement would not lessen the importance of it, nor would it invoke outrage or any other emotional response from other students or faculty present.

Finally, Joyce seems to overstate the breadth of moral disagreement present in society. It is rare to meet an individual with the view that murder is not wrong. In most situations, I would argue, we meet individuals who have the same, or very similar, moral standards, but merely disagree on how to achieve these standards. Olson directly opposes this view by arguing that in most nations, there is widespread moral disagreement. He draws attention to

the differences between socialists and neo-liberals, nationalists and cosmopolitans, and pro-life and pro-choice activists. This is a disagreement most people will see every day, he claims (2014, p. 127-128). However, I think Olson's argument misses the nature of moral disagreement. At a surface level, these people all appear to have competing moral standards, but realistically they probably share the same moral standards, but disagree on how best to achieve them. For instance, most people would agree that the preservation of human life is an important goal. One individual may believe that this can best be achieved by the implementation of socialist policies, while another could believe it is best achieved by neo-liberal policies. However, there is no fundamental moral disagreement present here. Even when a pro-life activist attempts to convince an individual with a pro-choice stance regarding abortion, the pro-life activist will usually attempt to appeal to the moral standards that they both have in common. Hence why they would claim "abortion is murder" because both parties would agree that murder is morally wrong, but merely differ on how murder manifests. Consequently, it does not seem that the question of whether moral values are relative or absolute would be something that would affect ordinary people, and so the goal of moral discourse would not be affected by the absence or presence of irreducible normativity (Finlay, 2008, p. 356).

Conclusion

Consequently, the standard error theory fails in proving that "irreducible normativity" is an essential feature of moral discourse. I analysed this claim in relation to Finlay's objections to the standard error theory and examined his reductionist account of moral value. I then considered Joyce's attempts to respond. Ultimately, I found that Joyce failed to prove that irreducible normativity is an essential feature of ordinary moral discourse, due to his misinterpretation of Finlay's account of relativism and his overstatement of the presence of moral disagreement in society. Thus, there is no issue with the "queer" relation of moral agents to moral facts, and we can reject the error theory on these grounds.

Bibliography

Anscombe, G.E.M., 1958. "Modern moral philosophy," *Philosophy*, 33: pp. 1–19.

Finlay, S. (2008) "The error in the error theory", *Australasian Journal of Philosophy*, 86(3), 347-369, DOI: 10.1080/00048400802001921.

Finlay, S. (2011) "Errors Upon Errors: A Reply to Joyce", *Australasian Journal of Philosophy*, 89(3), 535-547, DOI: 10.1080/00048402.2010.510531.

Foot, P. (1972) 'Morality as a system of hypothetical imperatives', *Philosophical Review*, 81(3), pp. 305–316.

Garner, R. (1990) "On the Genuine Queerness of Moral Properties and Facts," *Australasian Journal of Philosophy* 68(2), DOI: 10.1080/00048409012344161.

Joyce, R. (2003). *The Myth of Morality*. Cambridge: Cambridge University Press.

Joyce, R. (2011) "The Error In 'The Error In The Error Theory'", *Australasian Journal of Philosophy*, 89(3), 519-534, DOI: 10.1080/00048402.2010.484465

Lenman, J. (2013) "Ethics Without Errors", *Ratio*, 26(4), pp. 391–409. https://doi.org/10.1111/rati.12029.

Mackie, J. (1977). *Ethics: Inventing Right and Wrong.* London: Penguin Books.

Miller, A. (2013). *Contemporary Metaethics: An Introduction.* Cambridge: Polity Press.

Olson, J. (2014) *Moral Error Theory: History, Critique, Defence*. Oxford: Oxford University Press.

Robert Sparrow and Artificially Intelligent Warrior Robots

By Taylor Crates

As military technology continues to advance, arguments regarding the moral permissibility of certain weapon systems become more and more prevalent. This essay will evaluate Robert Sparrow's argument against the deployment of warrior robots on the grounds that no one can be held morally responsible for their actions. I will begin by defining the term 'warrior robots,' highlighting the possible benefits and ethical concerns brought about by their implementation. Secondly, I will offer a brief explanation of Robert Sparrow's argument and reveal why he believes that neither programmers, commanding officers, nor warrior robots themselves can be held morally responsible for the actions of a warrior robot. Thirdly, I will provide an overview of just war theory, elucidating the principles of *jus ad bellum* and *jus in bello*. Through the provision of relevant examples, I will clarify why the identification of morally responsible actors is a prerequisite for applying the principles of *jus in bello*. Fourthly, I will reject Sparrow's claim that neither the programmers nor commanding officers can be held morally responsible for the actions of warrior robots with reference to Jai Galliott's concept of shared responsibility. Finally, I will reveal that once warrior robots have been deployed, they gain the potential to disregard the principles of *jus in bello*. If this occurs, it will become unethical to continue their usage. Overall, morality as defined by just war theory does not forbid us from deploying warrior robots into battle so long as they function in a just manner.

Before approaching Robert Sparrow's argument, it is first necessary to define the term 'warrior robots' and acknowledge the potential benefits of using this technology in warfare. Warrior robots, while not in use currently, are predicted to be the next evolution of autonomous weapon technology (De Landa, 1991, p.170). Militaries around the world already utilise certain forms of autonomous weapon systems, including surveillance drones, torpedoes, and cruise missiles. In these examples, the word 'autonomous' simply refers to the fact that the devices can perform some part of their intended function without immediate human control. However, these systems still require some degree of guidance or assistance (Sparrow, 2007, p.63). Warrior robots, on the other hand, will be artificially intelligent enough to reason, make decisions for themselves, learn from their experiences, and act upon

their observations (De Landa, 1991, p.170). These systems will not only be able to function without human control but will be able to replace the role of human soldiers altogether. Deploying this technology has the possibility to provide a range of benefits (Meyers, 2011, p.22). For example, military objectives could be accomplished while subjecting fewer human beings to physical injuries, psychological damage, or the possibility of death. In addition to this, warrior robots are predicted to be able to fight in a more efficient and logical manner than human soldiers (Sparrow, 2007, p.64). However, these possible benefits do not mean that there are no objections to deploying warrior robots into battle. As these systems have the ability to learn from experience and act on their own, they also have the ability to become unpredictable (Sparrow, 2007, p.65). This prompts many questions regarding the ethicality of implementing such a technology.

Robert Sparrow serves as one of the many philosophers who have argued against the military's implementation of warrior robots. In his publication, "Killer Robots," he states that it would be unethical to use these systems in battle unless someone could be held morally (and possibly legally) responsible for the deaths and harm that they may cause. Specifically, he focuses on the possibility that these robots, due to their autonomy and unpredictability, could bring harm or death to innocent civilians (Sparrow, 2007, p.67). Such an atrocity would demand that someone be held responsible, whether it be the programmers of the robot, the commanding officers that deployed it, or the robot itself. However, so long as it is not a case of negligence, the programmers of the robot could not justly be held accountable for the wrongdoings of the robot itself. These artificially intelligent systems have the capacity to make their own decisions and function accordingly. Sparrow argues that holding the programmers responsible for the actions of an autonomous robot would be equivalent to holding parents responsible for the actions of their full-grown child (Sparrow, 2007, p.70). Sparrow continues to argue that it would also be unjust to hold the commanding officers responsible. Although orders from the commanding officers might have the ability to influence the robot's actions, the robot ultimately has control over its own decisions, including the decision of who to target (Sparrow, 2007, pp.70-71). Finally, Sparrow argues that the warrior robot itself cannot be held responsible for its own actions. Although the robots are autonomous to some degree, they are not capable of receiving blame or punishment as they do not meet the conditions of moral personhood (Sparrow, 2007, p.73). Sparrow concludes that, because no one can be held morally responsible for the deaths and damages caused by warrior robots, deploying these autonomous systems would result in the

war becoming unjust. Therefore, the utilisation of warrior robots in warfare is unethical (Sparrow, 2007, pp.74-75).

Sparrow's argument rests entirely on the claim that, in order for a war to be just, there must be someone who can be held responsible for any death that occurs as a result of said war. He states that the allocation of individual responsibility is a requirement, or precondition, necessary for applying the principles of *jus in bello* (Sparrow, 2007, p.67). Just war theory comprises a number of well-tested principles surrounding the morality of war, which are typically demarcated into two main categories (Meyers, 2011, p.23). The first category contains principles of *jus ad bellum*, which centre around **whether** it is morally permissible to go to war based on certain conditions (Lackey, 1989, p.29). For example, one of the most important principles of *jus ad bellum* is that the reason for fighting must be significant enough to justify the use of violent force (Lackey, 1989, p.33). If the commanding officers of a nation or political group choose to become involved in warfare, yet disregard this principle, the war would be unjust. Sparrow's argument, however, refers to the second category of just war theory, which contains principles of *jus in bello*.

The principles of *jus in bello* centre around **how** to engage in warfare in a way that is morally permissible (Lackey, 1989, p.29). There are essentially three primary principles within this category, which apply to both the actions of soldiers and the forms of weaponry employed. The first principle demands that all acts of war must be essential in accomplishing military objectives (Meyers, 2011, p.24). It is morally wrong for soldiers, or a style of weapon, to create more violence than is absolutely necessary. The second principle, or the 'proportionality' principle, requires that the importance of any military objective be significantly greater than the amount of violence and destruction it may cause. Finally, the 'discrimination' principle requires that only active military targets be attacked or harmed. Engaging in warfare in a way that inflicts violence on civilians, wounded soldiers, doctors, or surrendered troops is morally wrong (Meyers, 2011, p.25). Overall, if a war is conducted in a manner that disregards the principles of *jus in bello*, the war itself becomes unjust. Consider the following example:

A warrior robot is deployed by the South in the American Civil War. Before entering battle, the robot was programmed to fight in accordance with appropriate war conduct based on 'jus in bello' principles. Whilst in battle, the robot comes across a military doctor working for the

opposing side, the North. Although the doctor is not a legitimate target and carries no weapon of his own, the robot still decides to kill him. Having calculated the outcome, the robot purposefully prevented the doctor from tending to the wounded soldiers of the North.

As military doctors are not legitimate targets, it is clear that the warrior robot violates the discrimination principle of *jus in bello*. This act is morally impermissible. However, in order to properly apply these principles, it is necessary to identify the individual morally (and possibly legally) responsible for this action. In Sparrow's words, "It is a minimal expression of respect due to our enemy - if war is going to be governed by morality at all - that someone should accept responsibility, or be capable of being held responsible, for the decision to take their life. If we fail in this, we treat our enemy like vermin…" (Sparrow, 2007, p.67). If it is true that no one can be held morally responsible for the actions of a warrior robot, as Sparrow claims, then it is also true that the principles of *jus in bello* cannot be applied to their actions. This supports the argument that deploying warrior robots in warfare would be unethical, as it would be impossible to enforce that these autonomous systems fight in a way that is just (Sparrow, 2007, p.67).

As warrior robots are incapable of suffering or receiving punishment, it is understandable why one cannot justifiably hold these systems morally responsible for their actions (Sparrow, 2007, p.71). However, Sparrow's claim that both the programmers and commanding officers are exempt from moral responsibility seems entirely absurd. Although these parties may not be directly and solely responsible for the atrocious actions of a warrior robot, they are still at least *partially* responsible for failing to develop and deploy a technology that complies with universal principles of just warfare. Within his publication "No Hands or Many Hands? Deproblematizing the Case for Lethal and Autonomous Weapon Systems," Jai Galliott proposes the concept of shared responsibility, arguing that an individual does not have to hold complete control over an autonomous system in order to be morally responsible for its actions (Galliot, 2020, p.169).

Reconsider Sparrow's claim that holding the programmers responsible for the actions of a warrior robot would be equivalent to holding parents responsible for the actions of their adult children (Sparrow, 2007, p.70). It is true that parents do not hold complete control over the actions of their full-grown children. After reaching a certain age, 'children' are entirely independent and autonomous, just as any other human being. Similarly, programmers do not

hold complete control over the performance of warrior robots. As this technology is autonomous, the robots are capable of determining their own actions. However, Galliott argues that parents are at least *partially* responsible for the actions of their adult children as it was their responsibility to prepare them for independence (Galliot, 2020, p.170). Even though programmers are not entirely responsible for the actions of warrior robots, they are at least *partially* responsible as it was their initial responsibility to develop an autonomous system that would perform in a just manner. A similar analogy can be used to elucidate the role of commanding officers. Galliott argues that, when teaching their child to drive, parents have a responsibility to conduct this lesson in a space where the child can learn safely without putting anyone's life at risk (Galliot, 2020, p.170). Similarly, commanding officers have a responsibility to thoroughly test warrior robots in a safe environment, ensuring their probability of acting in a just manner before deploying them in battle. In the case that a warrior robot commits a war crime or some other atrocity, the commanding officers who deployed it are at least *partially* responsible, as it was their responsibility to test and guarantee the safety of these autonomous systems before using them. Unlike Sparrow, Galliott takes into consideration the complexity of this technology and denies the need for an individual to hold full moral responsibility (Galliot, 2020, p.171).

As for the principles of *jus in bello*, having the potential to allocate moral responsibility is a precondition for the application of these principles (Sparrow, 2007, p.70). However, this moral responsibility does not have to be assigned to one individual. If we reconsider the example on page five with an understanding of Galliott's concept of shared responsibility, it is possible to hold multiple individuals *partially* responsible for the morally impermissible act of the warrior robot. The programmers of the robot can be held *partially* morally responsible as they programmed an autonomous system capable of disregarding the discrimination principle. The commanding officers of the robot can also be held *partially* morally responsible as they deployed an autonomous system without first ensuring that it would obey the discrimination principle. The concept of shared responsibility for warrior robots reveals that, contrary to Sparrow's argument, it is possible to justifiably hold someone (or someone's) morally responsible for the actions of an autonomous warrior robot (Sparrow, 2007, pp.74-75). Therefore, the responsibility precondition for applying the principles of *jus in bello* is met as well (Sparrow, 2007, p.67).

Given that warrior robots are intended to function in a morally permissible manner, it does not counteract the principles of just war theory to simply deploy them in battle (Sparrow, 2007, p.69). However, this does not guarantee that their usage will remain morally permissible once deployed. As explained, the principle of discrimination applies to both the actions of soldiers and the forms of weaponry employed (Meyers, 2011, p.25). Unfortunately, the harming of illegitimate targets is a common occurrence in most acts of war. This can be related back to the principle of proportionality. So long as the harm brought to illegitimate targets is minor in comparison to the overall importance of the military objective, a small oversight of the principle of discrimination is permissible (Lackey, 1989, p.60). For example, if a warrior robot were to bring harm to a single military doctor, this act would be considered morally impermissible, and the programmers and commanding officers of the robot would share moral responsibility for this action (Galliot, 2020, p.171). However, this action alone is not enough to deem warrior robots an unjust form of weaponry. But, if this act becomes a relatively common occurrence, and warrior robots reveal an overall inability to discriminate, continuing to use these weapons in battle would become unjust. In 1997, an international convention met to discuss the use of anti-personnel mines in acts of war. The convention concluded that these mines should be prohibited based on their inability to discriminate between combatants and civilians, "causing unnecessary suffering or superfluous injury" (ICRC, 2021). Similarly, if warrior robots as a whole prove unable to comply with the principle of discrimination of *jus in bello*, it will become morally necessary to ban their usage.

In conclusion, the potential development of warrior robots as a form of autonomous weaponry brings to question whether it will be ethical to use these systems in warfare. Within his argument, Robert Sparrow illuminates some of the potential dangers brought about by these systems, arguing that morality forbids their deployment. Although no one is in direct control over the actions of warrior robots, Sparrow fails to recognize that both the programmers of the robots and the commanding officers who deployed them can be held *partially* morally responsible for any atrocities they may commit. So long as these autonomous weapons function as intended, the principles of *jus in bello* do not prohibit their deployment. However, if warrior robots prove to disregard the principles of just war theory once implemented, morality would forbid us from continuing to use them.

Bibliography

De Landa, M., (1991). *War in the Age of Intelligent Machines*. New York: Zone Books.

Galliott, J., (2020). No Hands or Many Hands? Deproblematizing the Case for Lethal
   Autonomous Weapon Systems. In: Roach, S.C. and Eckert, A.E., eds. *Moral*
   Responsibility in Twenty-First-Century Warfare: Just War Theory and the Ethical
   Challenges of Autonomous Weapon Systems. Albany: State University of New York
   Press, pp.155-180.

ICRC, (2021). *1977 Anti-Personnel Mine Ban Convention*. Geneva: International Committee
   of the Red Cross.

Lackey, D.P., (1989). *The Ethics of War and Peace*. Englewood Cliffs: Prentice Hall.

Meyers, C.D., (2011). GI, Robot: The Ethics of Using Robots in Combat. *Public Affairs*
   Quarterly, 25(1), pp.21-36.

Sparrow, R., (2007). Killer Robots. *Journal of Applied Philosophy*, 24(1), pp.62-67.

De Landa, Manuel. (1991). *War in the Age of Intelligent Machines*. New York: Zone Books,

Galliott, Jai. "No Hands or Many Hands? Deproblematizing the Case for Lethal Autonomous
   Weapon Systems." In *Moral Responsibility in Twenty-First-Century Warfare: Just*
   War Theory and the Ethical Challenges of Autonomous Weapon Systems. Edited by
   Steven C. Roach and Amy E. Eckert. Albany: State University of New York Press,
   2020. 155-180.

ICRC. (2021). *1977 Anti-Personnel Mine Ban Convention*. Geneva: International Committee
   of the Red Cross,

Lackey, Douglas P. (1989). *The Ethics of War and Peace*. Englewood Cliffs: Prentice Hall,

Meyers, C. D. (2011). "GI, Robot: The Ethics of Using Robots in Combat." *Public Affairs Quarterly* 25, no. 1 (2011): 21-36.

Sparrow, Robert. (2007). "Killer Robots." *Journal of Applied Philosophy* 24, no. 1 (2007): 62-67.

On the Role of Absences in Causal Explanation

By Rían Coady

In this paper, I argue that absences are not causes but can feature in causal explanations. I first argue, following Beebee, that although we intuitively count some absences as causes, there are also many more that we do not, and the friend of causation by absence cannot draw a principled metaphysical distinction between those absences that cause and those that do not. Thus, the intuition that absences are causes is insufficient to overturn the standard picture of causation as a relation between events. However, absences can still be given a role in causal explanation, and common sense's error in counting some absences as causes can be explained by its failure to distinguish between causation and causal explanation and between true and adequate explanations. I then consider an objection to Beebee's account from Mumford, who argues that explanations featuring absences cannot be causal. I argue that Mumford misrepresents Beebee in suggesting that absences must be *part of* an event's causal history when they need only provide information *about* the event's causal history. I then outline Mumford's account of assertion and denial and suggest that we can appeal to this to avoid any negative facts on Beebee's account. Finally, I argue that Beebee's view is to be preferred to Mumford's because it is more plausible that common sense confuses causal explanation—rather than counterfactuals—with causation.

Intuitively, it seems that absences can be causes; e.g., it seems that the gardener's failure to water the plants caused them to die. Absences, however, are not events—they are nothings—which poses a problem for the standard picture of causation which considers causation to be a relation between events. Beebee calls this the 'relationist' view, in opposition to 'nonrelationist' views, which give up this picture. Although common sense does count some absences as causes, there are many more that it does not, and as we shall see, the nonrelationist has no principled way of distinguishing between them (Beebee, 2004). A first attempt at analysing causation by absence might go as follows (Beebee, 2004, p. 294):

(I) The nonoccurrence of an event of type *A* caused event *b* if, and only if, had an
   A-type event occurred, *b* would not have occurred.

This definition agrees with our intuitions about the gardener's failure to water the plants being the cause of their death; if she had watered them, then they would not have died. However, this analysis also counts as causes of all sorts of absences that common sense does not recognise as causes, e.g., Boris Johnson's failure to water the plants (Beebee, 2004).

Beebee (2004) suggests that common sense counts the gardener's failure as a cause, but not Johnson's, because the gardener *usually* waters the plants, and her failure to do so was an *abnormality*, whereas Johnson never waters them. Beebee notes that we also intuitively count as causes those absences which violate norms. For example, if a dog gets an eye infection and goes blind, we intuitively say that the dog owner's failure to get the infection treated was a cause of the dog's blindness, even if his mistreatment of the dog is not abnormal. Here, what makes common sense is to count the dog owner's failure as a cause, but not Johnson's failure, because the dog owner has a particular moral responsibility to look after his pet, whereas Johnson has no such obligation. Legal or professional norms may also inform our intuitive causal judgments; e.g., even if the gardener is lazy and frequently fails to do her job, we intuitively still count her as the cause of the plants' death since she is shirking her professional duties. The same goes for epistemic norms: common sense might not count a drug company's failure to disclose the side effects of their drug as a cause of people getting ill if the side effects were unforeseeable, but if the company knew about it in advance, or the effects were reasonably foreseeable, then we intuitively count their failure as a cause. These considerations suggest a refined analysis of causation by absence (Beebee, 2004, p. 296):

(II)   The absence of an *A*-type event caused *b* if and only if
(i) *b* counterfactually depends on the absence: Had an *A*-type event occurred, *b* would not have occurred; and
(ii) the absence of an *A*-type event is *either* abnormal *or* violates some moral, legal, epistemic, or other norm.

Now, as Beebee says, this looks like a fine analysis of our *concept* of causation by absence but is not something we should be providing as a metaphysical account of causation by absence. If we take (II) to give the *truth conditions* for claims about causation by absence, then their truth will depend on normative factors. But this is absurd—the causal facts do not depend on professional gardening norms! It also makes causation a relative matter in that if there is disagreement over, say, what is reasonably foreseeable, then there will also be disagreement over whether the drug company caused the side effects or not. In addition to

these issues arising from normativity, causal claims will also become hopelessly vague because of the reference to abnormality. There is no clear boundary between what is normal and abnormal (how many times does Johnson have to water the plants before that becomes 'the new normal'?), so causal claims will be indeterminate, which is also unacceptable (Beebee, 2004).

The final analysis Beebee (2004, p. 298) considers is:

(III)   The absence of an *A*-type event caused *b* if and only if
(i) if an *A*-type event *had* occurred, *b* would not have occurred; and
(ii) an *A*-type event occurs at a world that is *reasonably close* to the actual world.

(III) seems to rule out many spurious cases of causation without bringing abnormality and normativity into play. It still rules that the gardener's failure to water the plants caused them to die, while Johnson's failure did not because the nearest world where the gardener watered them is much closer to actuality than the nearest world where Johnson did so. Through this analysis, abnormality and norm violations are not part of the truth conditions of causal claims but merely inform us how close those worlds are where the absence occurs. While this is an improvement over the previous analyses, it still fails, since in the case of the cruel dog owner, even though the nearest world where he treats his dog is very distant from ours, we intuitively still judge his failure to be the cause of the dog's blindness. It seems like it is really the norms that are doing the work here and directly deciding our intuitive causal judgments and not merely informing our judgments about the closeness of worlds. (III) also suffers from the same indeterminacy issue as (II). Even if we can order worlds in terms of their *relative* similarity to actuality, what counts as 'reasonably close' is as vague as what counts as 'abnormal' (Beebee, 2004).

Beebee (2004) thus shows that all of the proposed nonrelationist analyses of causation by absence fail for one reason or another. Thus, although the standard relationist view will have to go against intuition in saying that no absences are causes, this is not any greater a price than one has to pay on the nonrelationist view anyways; either one also has to pay the price of going against intuition, or one ends up with an analysis of causation that makes causation normative, relative, vague, or indeterminate. Hence, the intuition that some absences are causes does not give us sufficient reason to reject relationism (Beebee, 2004).

However, Beebee (2004) suggests we can soften the counterintuitive blow of denying that absences are causes by holding that absences can still be given a role, not in causation,

but in *causal explanation*, and we can also explain why intuition goes wrong by seeing how it confuses the two. She draws on Lewis's (1986) account of causal explanation, according to which 'to explain an event is to provide some information about its causal history' (p. 217). Beebee (2004) distinguishes between causation and causal explanation as follows:

the canonical form of causal statements is "*c* caused *e*," where *c* and *e* are events and *caused* is a two-place relation, rather than "*E* because *C*," where *C* and *E* are facts and "because" is a sentential connective. (p. 301)

To see that causation and causal explanation can come apart, consider the following example (Beebee, 2004, p. 302):

(1) Oswald's shot caused JFK's death.
(2) JFK died because Oswald shot him.
(3) JFK died because somebody shot him.

(1) is a causal claim; it says that one event caused another. (2) is an explanation, and clearly, a causal one since the explanans is the fact that the cause happened and the explanandum is the fact that the effect happened. (3) is also a causal explanation, but not as obviously so; here, the explanans does not stand to the explanandum as cause to effect, because there is no such event as *someone's* shooting JFK. This would be a bizarre disjunctive event consisting of Oswald's shooting him, or my shooting him, or Boris Johnson shooting him, and so on. But 'there is no more such an event than there is an event of my-birthday-party-or-your-morning-bath' (Beebee, 2004, p. 303). Nevertheless, (3) is still a causal explanation because we are given some information about the effect's causal history, viz., that it included some event, that was a shooting of JFK by someone. Thus, we can have causal explanation without causation, and Beebee suggests that we should treat absences the same way. Even if the gardener's failure to water the plants did not cause their death, we could still say that the plants died because the gardener failed to water them. This tells us something about the death's causal history, viz., that it did not include an event of the gardener's watering the plants. Thus, even if they are not causes, absences can still be given a role in causal explanation. We also see from this why common sense makes the mistake of counting absences as causes. In ordinary cases of causation, causation and causal explanation go together, so in most cases, we can infer from '*E* because *C*' to '*c* caused *e*', but common sense goes wrong in applying this principle to absences (Beebee, 2004).

Beebee (2004) suggests that we can also explain why common sense counts some absences as causes but not others by considering the distinction between *true* and *adequate* explanations. Strictly speaking, 'The plants died because the gardener failed to water them' and 'The plants died because Johnson failed to water them' are equally true—they both give us the correct information about the death's causal history, viz., that it did not include an event either of the gardener or of Johnson watering the plants—but only the former is adequate. Adequate explanations must be sensitive to why the explanation was requested in the first place. For example, someone looking to know why the plants died is probably looking for someone to blame, in which case it would be correct to point the finger at the gardener but totally misleading to mention Johnson. Again, this distinction between true and adequate explanations is one that common sense is likely to ignore, so we can further explain how its errors arise (Beebee, 2004).

Mumford (2021) agrees with Beebee that absences can be explanatory without being causes but disagrees that the explanations involved are *causal* explanations:

> *I cannot see how absences, if they really are nothing at all, can be part of something's causal history since absences cannot be either causes or causally relevant. Nor would I be happy including negative facts in a causal history, as an intended less innocuous alternative to absences… Until such time as a credible account is found for how absences, or any other negative elements, play any causal role at all, then we should resist the idea that they are part of the causal history. (p. 83)*

Instead, Mumford thinks that the explanatory power of absences comes from their featuring in true counterfactuals about what would have happened if they had been present; e.g., 'The plants died because the gardener did not water them' is explanatory because, had the gardener watered the plants, they would have lived. Similarly to Beebee, he thinks that pragmatic and normative considerations play a role in selecting the relevant counterfactual and making some explanations more appropriate to assert (Mumford, 2021).

However, Mumford has misrepresented Beebee's account of absences' explanatory role. He seems to think that Beebee is suggesting that absences are *part of* an event's causal history (he uses this phrasing twice in the above quote) and that absences are part of the causal history in the same way that the positive causes of that event are, such that if we were to represent an event's causal history by a list of all the events that caused it, we would list both the presences that actually did the causing, but also a bunch of absences that did not

cause it. If this were what Beebee was saying, Mumford would be right to object that this is giving absences a problematic metaphysical status. However, he has misrepresented Beebee here. Beebee is not saying that absences show up anywhere on the list, so to speak; instead, she is saying that when we have an absence in a causal explanation, we are providing information *about the list* (causal history), viz., what is *not* written on it (that the causal history *does not* contain such-and-such events). Consider an analogy: if you ask who attended the last Metaphysics tutorial, and I tell you that Ciara did not attend, then I am not saying that the roll for the tutorial must look like this:

Present
Áine
Brendan
…
Absent
Ciara
…

The roll can consist just of the Present list without an additional Absent list. Rather, when I say that Ciara was absent, I am giving you information *about the roll*, viz., that Ciara's name is not on it. But I am not thereby including any absences in the tutorial history. Similarly then, Beebee's account is not committed to including absences in events' causal histories, yet they can still play a role in causal explanation by giving us information about an event's causal history.

In the above Mumford quote, we see him resisting the idea of negative facts being included in an event's causal history. He might argue that Beebee's account is still committed to problematic negative facts and that these are just as bad as absences. Here, however, we can appeal to Mumford's own account of assertion and denial to avoid this. Breaking with a tradition that goes back to Frege (1960), Mumford (2021) argues that denying P is not equivalent to asserting not-P, based on the distinct functions of assertion and denial. On his view, denying, say, that there is a hippo in the room is not equivalent to asserting a negative fact of there not being a hippo in the room. Applying this to absences, we need not be committed to any negative facts showing up in events' causal histories when we say absences provide information about what does not show up in an event's causal history. Instead, when I say, 'The plants died because the gardener did not water them', I am simply denying that there was an event of the gardener's watering in the causal history of the death—no negative fact needed.

One may wonder why one would bother to go to such lengths to defend Beebee's view against Mumford's, given their major similarities and their agreement on the essential points. To me, a major appeal of Beebee's view is that, by counting the explanations that absences feature in as *causal* explanations, she goes a long way in explaining why common sense goes wrong in counting absences as causes. On Mumford's view, the 'because' involved in the explanations that absences feature in is a non-causal 'because', which increases the distance between causation and the kind of explanations involved, making it a slightly less satisfying account of how common sense goes wrong. Although there is a close connection between causation and counterfactuals, the connection between causation and causal explanation is that bit closer, and I think that the view that says that common sense mistakes something more similar to causation (viz., causal explanation) than something slightly less similar (viz., counterfactuals) is to be preferred. Of course, this is only a slight reason to prefer Beebee's view, but when we are comparing views this similar, it is to be expected that the benefits of one over the other will be this slight. Moreover, a slight benefit is still a benefit, and enough for me to lean towards Beebee rather than Mumford.

This paper argued that absences are not causes but can feature in causal explanations. It outlined the relationist view of causation and showed that it is no more in danger from the intuition that some absences are causes than its nonrelationist counterpart; there are many more absences that common sense does not count as causes, and the nonrelationist must either pay a similar price in going against intuition, or else ends up with causal claims becoming normative, relative, vague, or indeterminate. Nevertheless, we can still give absences a role in causal explanation and can account for common sense's error by showing how it confuses causation with causal explanation and true explanations with adequate ones. Mumford's objection to Beebee's account rests on a misunderstanding, and we see that Beebee's account need not hold that absences are part of causal histories but need only provide information about them. This does not commit us to any negative facts since we can appeal to Mumford's distinction between assertion and denial. Beebee's view is slightly preferable to Mumford's because hers attributes less error to our intuitive causal judgments since causal explanation is closer to causation than counterfactuals are.

Bibliography

Beebee, H. (2004). Causing and Nothingness. In J. Collins, N. Hall, & L. A. Paul (Eds.), *Causation and Counterfactuals* (pp. 291–308). MIT Press.

Frege, G. (1960). Negation. In P. Geach & M. Black (Eds.), *Translations from the Philosophical Writings of Gottlob Frege* (pp. 117–135). Blackwell.

Lewis, D. (1986). Causal Explanation, in his *Philosophical Papers* (Vol. II, pp. 214–240). Oxford University Press.

Mumford, S. (2021). *Absence and Nothing: The Philosophy of What There is Not.* Oxford University Press.

A Critique of Priest's Master Argument for Dialetheism

By Darragh Senchyna

Graham Priest, in both (Priest, 1996) and again in (Priest, 2001), developed an argument for the existence of true contradictions (dialetheia), which he derived from Berkeley's master argument for idealism. Priest's argument has two forms and relies on several lemmas. The thesis of the following paper is that both forms of argument fail but for different reasons. The former fails due to a lack of consideration of the temporal aspect of conceivability, and the latter has issues with intelligibility and allows for much less radical alternatives to dialetheism. To begin, I will introduce Berkeley's master argument for idealism and analyse a particular formalisation of the argument. I will then offer some preliminary analysis of self-refutation arguments generally before introducing Priest's argument as he characterised it in Priest (2001), along with the lemmas needed to make the argument valid. I will then demonstrate the issues with both forms of Priest's argument and conclude that it is unsuccessful in arguing for true contradictions.

What I call[1] Berkeley's master argument for idealism is an argument George Berkeley first articulated in *A Treatise Concerning the Principles of Human Knowledge*, which also appears prominently in the *Three Dialogues Between Hylas and Philonous*. For this essay, I will look at a simplified version of the argument drawn from the latter of these two texts. The argument there deals with a puzzle demonstrated by a conversation between two characters in the dialogues, Hylas and Philonous.

Philonous asks Hylas if he can conceive it possible for any sensible object to exist without it being conceived[2]. Hylas says that he can do this since he can conceive of a tree or a house existing without any minds conceiving it. Philonous then interjects that the tree and the house Hylas mentioned were conceived by Hylas' mind and must be conceived in order for

---

[1] And what is usually called see: (Prior, 1955), (Mackie, 1965), (Priest, 1996), (Priest, 2001), (Levine, 2013)

[2] Berkeley does not use this term, and the characterisation above likely does not do justice to Berkeley's true intentions. I ask for leniency here since Berkeley's argument itself is not what is principally discussed but Priest's augmentation of it.

him to use them as examples. Hylas concedes that he cannot conceive of any sensible object unconceived by a mind. The attempt is self-refuting.

What does this argument establish? The deduction begins with Philonous asking Hylas if he can show him through example that anything can exist unconceived, and it ends with the conclusion that Hylas cannot do this. In more formal terminology, we may say that Philonous claims that:

(I) $\neg(\exists x \neg \tau x)$

Where τ means 'is conceived' or 'is thought'. Philonous then asks Hylas to prove him wrong by example, to devise a 'c' such that τc. Hylas cannot do this, and it is seemingly impossible. There are multiple ways of formalising this argument. Below, I will attempt to do so in a manner that integrates nicely with our planned discussion of Priest but also owes some debt to Prior (1955).

We should first distinguish between two syntactically different forms of conception involved in the argument. There is the conception of a particular object, 'c', and there is the conception of a thought or sentence, 'c is unconceived'. We will use 'τ' for the former and 'T' for the latter. The first lemma we need to accept is that conceiving a sentence which involves an object involves conceiving that object.[3] In other words,

$$T_{\varphi(c)} \Rightarrow \tau c$$

Where $\varphi(c)$ is a sentence which involves the object c. The argument then proceeds in the following way (assume that stating/naming implies conceiving):

---

[3] Prior avoids introducing this lemma by requiring an agent for T and explicitly defining it.

(1.0)	(Hylas states that something exists unconceived)
(1.1) $\;^{T}{}_{\Rightarrow \exists x T\neg\tau x}$	(Hylas names a thing which he says is unconceived)
(1.2) $\Rightarrow \dfrac{\neg\tau c}{\tau c}$	(Hylas conceived that thing)
(C) $\Rightarrow \tau c \wedge {}^{T}{}_{\neg\tau c}$	(Hylas conceived a thing and conceived it as being unconceived)

First, note that the argument we have now characterised does not result in a true contradiction. (C) states that both c is conceived, and c is conceived of as being unconceived. In Prior's terms, the argument shows that $\neg\tau c$ cannot be conceived *truly*. It cannot be both true and conceived. Secondly, the inference from (1.0) to (1.1) is not immediately valid. We have not stipulated that conceiving of there existing a thing *generally* implies conceiving of some particular thing. If we begin with (1.1) as an assumption, then the argument only proves that it is not possible to name an unconceived thing, not that it is impossible to conceive that there generally exists an unconceived thing.

The meaningful part of the argument lies in a question which is more difficult to formalise. Does it make any sense to generalise over things which principally cannot be named? It could be argued that we do so regularly. We can name things which were unconceived five minutes ago. In fact, when we include a temporal element then the take-aways from Berkeley's argument become less determinate. Could it not be the case that c was unconceived at (1.1) but after (1.1) became conceived? Or did it all happen at the very same instant of time?

One possible solution would be to change c from 'c is conceived' to 'c is conceivable'. Once we do this, we lose the possibility of concluding anything similar to Berkeley's idealism, but we get greater insight into the issue that his argument illuminates. In some sense, the question of whether it is possible to generalise over things which it is self-refuting to pick out individually is dealt with in Davidson (1973). Davidson claims here that it is meaningless to postulate about conceptual schemes which are not translatable into our own. He determines this by arguing that it is self-refuting and thus impossible for us ever to gain any knowledge of these conceptual schemes. The historical and contemporary debate around this form of

argument is dealt with further in Levine (2013), but for this paper, we only need to note that the possibility of generalising over inconceivable things is a contentious issue.

Priest, of course, is not using Berkeley's master argument to make claims about generality but to make claims about true contradictions. His formalisation in both (Priest, 1996) and (Priest, 2001) looks like the following.

(1.0) $T\exists x \neg \tau x \ \neg \tau c$

(1.1) $=> T$

(1.2) $=> \tau c$

(2.0) $\exists x \neg \tau x$

(2.1) $=> \neg \tau c$

$$(C) => \tau c \land \neg \tau c$$

Priest's argument looks a lot like ours but begins with two assumptions rather than one; it is conceived that something is unconceived, and additionally, there *is* something which is unconceived. He concludes with the conjunction that c is conceived and c is not conceived as a true contradiction. If Priest's argument is sound, we seem left with three options. Either Berkeley is correct, and we must reject (2.0), meaning we must adopt Berkelian idealism. Alternatively, the negation of (1.0) holds, and as Priest puts it, "*(Berkeley's) opponents cannot even conceive of their argument*" (Priest, 2001). Finally, the third option, which is what Priest is arguing for in this interpretation: Both (1.0) and (2.0) are true and result in a 'true contradiction at the limits of thought'.

Of course, the argument as it stands is not at all immediately valid. All of the inferences except the final one to the conclusion (C) need explanation, and Priest provides these explanations. We have already given a lemma for the inference from (1.1) to (1.2); Priest calls this lemma the 'Conception Schema'.

The next inference to deal with in Priest's argument is (2.0) to (2.1). To explain this inference, Priest uses what he calls 'Hilbert's ε-schema'. This refers to Hilbert's ε-terms, terms

popularised by David Hilbert for use in mathematical logic (Leisenring, 1969). Priest has his own formalisation of how terms like this might work in ordinary language (Priest, 1979).

Essentially, 'εxφx' is a term which arbitrarily chooses an object which satisfies φ if there is one. Priest says that this may be done through an 'official' choice function, functions which he attempts to capture in Priest (1979). The inference also relies on 'Hilbert's ε-axiom' which is the following:

$$\exists x\varphi x \rightarrow \varphi(\varepsilon^x\varphi^x)$$

Essentially, if there is an x that satisfies the property φ, then 'εxφx' satisfies that property. Priest then states that c = εx¬τx. Hence, (2.0) to (2.1). Finally, there is (1.0) to (1.1).

In order to establish this, Priest uses a combination of his assumptions around the ε -schema and what he calls the 'Prefixing Principle'. The Prefixing Principle states that conceiving 'prefixes' as logical consequences, that is:

If |- α → β then |- T(α) → T(β)

In other words, to conceive that A is also to conceive of all its logical consequences. This principle seems rather extreme. Priest recognises this but emphasises the limited scope of the principle in the argument above, stating it *"seems quite harmless in the present context precisely because Philonous is clearly taking Hylas through these consequences - or the relevant ones anyway."* (Priest, 2001).

The Prefixing Principle along with the ε-schema then allows us to make the inference from (1.0) to (1.1) and with that, fully explain Priest's argument. Clearly, the argument heavily relies on these lemmas. However, the lemmas are difficult to dismantle individually, and the problems with this argument, like the argument itself, are subtle.

The main issue with the argument comes primarily from how the Prefixing Principle interacts with Priest's definition of c. Priest states that an 'official' choice function could choose the 'c' in the above deduction. Let us call π a function of this kind. So $\pi(\neg\tau) = c$. The purpose of characterising the function in this way is purely for illustration. Functions of this kind involve many more variables (Priest, 1979). For our purposes, π is used to show that

there is a function at work in deciding c and that it depends on $\neg\tau$.

Now let us add two extra premises to Priest's deduction and look at the steps of the argument line-by-line rather than dividing it into two sides.

(1.0) $T\exists x\neg\tau x$

             $\neg\tau(\pi \; \neg\tau$

(1.0.1) => T   (   ))
(1.1)   => $T\neg\tau c$

(1.2)   => τc

(2.0) $\exists x\neg\tau x$

             $\neg\tau(\pi \; \neg\tau$

(2.0.1) =>   (   ))

30

(2.1)   => ¬τ*c*

(C) => τ*c* ∧ ¬τ*c*

  Let us walk through the argument. We begin with (1.0), the assumption that it is conceived that there is something which is not conceived. Then we use our choice function and conclude that the object 'c' is conceived. Next, we assume (2.0) that there is something which is not conceived. We once again utilise our choice function to choose such an object, and our choice function once again chooses c. However, this does not seem right. We have already established that c is conceived. Unless we have previously assumed that what we are conceiving[4] of could be dialetheic, this choice function now seems to violate Hilbert's ε-axiom. The choice function depends on the extension of ¬τ, but the extension of ¬τ

changes throughout the argument. In particular, before (1.1), it includes 'c', and afterwards, it does not. The issue here is with the Prefixing Principle. The principle states that to conceive that A is also to conceive of any of its logical consequences. However, surely it can not be the case to conceive that A is also to *simultaneously* conceive of all of its logical consequences. Priest even alludes to the temporal nature of the principal when he justifies it because Hylas is being 'walked through' the relevant consequences.

  Thus, the Prefixing Principle must be reformulated to include a temporal element. Such as;

  If |- α → β then |- $T_p(\alpha) \to T_q(\beta)$ where p < q

---

[4] This problem does not appear so readily when we walk through the argument in a different order, but that should not invalidate what we discover.

Which can be interpreted as, if A is conceived at a time then a logical consequence of A is conceived at a later time. The deduction no longer works if we add these time subscripts to T and τ. At best, we end up with the conjunction $\tau_p c \wedge \neg \tau_q c$; that is, c was conceived at time p and was not conceived at a later time q, which is not a contradiction. Of course, we already foresaw this issue and decided to move towards a form of the argument where τc

means 'c is conceivable' rather than 'c is conceived'. Priest uses this same move as a fail-safe against counterarguments of this kind. In this case, the Prefixing Principle seems much more plausible; If something is conceivable then so are all of its logical consequences. However, with this move comes two new problems for Priest.

First, there is now an apparent incoherence in using the choice function in the move from (1.0) to (1.1). How could we pick out some particular thing which is inconceivable to us? In a sense, this impossibility was precisely what we had previously determined this form of the argument proved. Priest argues that we could think of 'c' as being picked out *de dicto.* In other words, we should think of 'c' as just being the thing which is named from '$\varepsilon x \neg \tau x$' with no interpretation of it as an extensional thing or as having any unique properties other than that. However, even if we accept that such a choice function could actually designate the same object on both sides of the deduction, it seems relatively unclear how we could get epistemological access to a choice function like this, particularly when we have no knowledge and cannot get any knowledge of the set it is choosing from. The argument then requires that we at least know that the set of inconceivable things is non-empty. In fact, this knowledge is premise (2.0) of the deduction.

At the end of the section on this argument in Priest (2001), he recognises that the knowledge of the set of inconceivable things being non-empty is essential to his argument and states the following to back up that assumption:

"Continua being dense, there is no way that any particular point can be picked out by our finite perception; and, since there is an uncountable number of points but only a countable

number of descriptions in our language, there is no way that every point can be picked out by a description." (Priest, 2001)

In other words, some points on a continuum, presuming it is dense, are inconceivable to us since we have no perceptual or descriptive methods to pick them out as particulars. So inconceivability is defined here by the absence of a function which designates the object as a particular. However, is the choice function Priest wishes to use in his argument, not such a function? By design, it designates objects properly enough for us to deduce that they are conceivable. Priest has made conceivability such a weak concept through legitimising (1.0) to (1.1) that there no longer seems to be a non-circular way of demonstrating our knowledge of the existence of inconceivable objects.

Secondly, even if we could accept (1.0) to (1.1) and that 'c' on both sides referred to the same thing, the resulting contradiction no longer presents us in this form with a choice between Berklean idealism and dialetheia. Rejecting (1.0) is compatible with much less extreme philosophical positions. In particular, it means adhering to philosophical positions akin to that of Davidson (1973), as we mentioned earlier. In fact, the incoherence of the premise that one can conceive of there existing something inconceivable is exactly how we characterised an interpretation of Davidson's position.

To conclude, Berkeley's master argument for idealism can be formalised and augmented to elucidate interesting issues around conceivability and thought. However, Priest's specific augmentation of it to argue for true contradictions at the limits of thought fails. In the first instance, where we consider c to mean 'c is conceived', it fails because the Prefixing Principle does not take *time* into account. In the second instance, where c is interpreted to mean 'c is conceivable', it fails partially due to issues around the existence of a choice function but also since the argument in this form has many more appealing alternatives. If Priest means to use this argument to convince readers of a conclusion so unorthodox as the existence of true contradictions, his argument should not be so permeable.

Bibliography

Davidson, D. (1973). On the Very Idea of a Conceptual Scheme. *Proceedings and Addresses of the American Philosophical Association*, 47, 5.

Leisenring, A. C. (1969). Mathematical Logic and Hilbert's & Symbol. *London: Macdonald Technical & Scientific.*

Levine, J, (2013). 'Logic and solipsism', in Peter Sullivan, and Michael Potter (eds), *Wittgenstein's Tractatus: History and Interpretation*, Mind: 9

Mackie, J. L. (1965). Self-Refutation--A Formal Analysis. *Journal of Symbolic Logic* 30 (3):365-366.

Priest, G. (1979). 'Indefinite descriptions'. *Logique Et Analyse* 22 (85):5.

Priest, G. (1996) 'Some Priorities of Berkeley'. *Logic and Reality Essays on the Legacy of Arthur Prior*

Priest. G (2001) The Limits of Conception. *Beyond the Limits of Thought.* Oxford University Press; 2nd edition: 56-70

Prior, A. N. (1955). Berkeley in logical form. *Theoria* 21 (2-3):117-122.

The 'Darwinian Dilemma' Revisited: Clarifying the Site of Disagreement between Street and Copp

By Frank Wolfe

Introduction

Moral realism is the view that there are moral facts, which are mind-independent. An important task for realist theories of morality is to explain why their purported moral facts have similar content to our common-sense moral beliefs. If they cannot explain this correlation, they are committed to the highly unappealing conclusion that many or all of our moral beliefs are simply factually incorrect, which undermines much of moral realism's apparent intuitiveness. In her seminal paper 'A Darwinian Dilemma for Realist Theories of Value' (2006), Street argued that any attempts by moral realists to explain this correlation run into a dilemma when we consider the findings of modern moral psychology. Given that our moral beliefs have been influenced by 'evolutionary forces' (because the beliefs which best promoted survival are the ones which persisted), realists have to either A) endorse the improbable claim that our current post-evolution moral beliefs just happen to match the moral facts by pure coincidence, or B) defend the scientifically implausible theory that moral facts have causally influenced our moral beliefs. Copp (2008) suggested that his 'Society-Centred Theory' of natural moral realism was immune to this dilemma. Street (2008) vehemently denied this, instead, accusing Copp of 'trivially restating' the first horn.

In this essay, I submit that Street is ultimately correct to argue that Copp's approach fails to escape this problem, but Street's case is more convincing if we clarify the site of disagreement between her and Copp. Specifically, I reformulate Street's challenge as a Darwinian *Trilemma*, which more charitably explains how Copp thinks his approach avoids the initial dilemma, and more accurately demonstrates why his approach fails. First I explain what Street means by 'evolutionary forces' and how she mobilises this into an argument against realism. Then I set out the disagreement between Copp and Street, and argue that Street's characterisation of Copp's approach is uncharitable. Finally I offer an alternative picture of Street's argument (the admittedly less catchy Darwinian *Trilemma*) and explain why this is even more damning for Copp's realism.

Street's Darwinian Dilemma

At heart, Street's Darwinian Dilemma questions the moral realists' ability to explain why our moral beliefs and the moral facts have similar or identical content. This challenge is made sharper by the fact that we seemingly have good evidence that evolutionary processes have pushed our ancestors' moral beliefs in certain directions. If moral realists cannot explain why the moral facts have gone in that direction too, they are forced to concede the 'implausible skeptical result' that most or all of our moral judgements are 'off track' (Street, 2006, p. 109).

Our first task is to understand why evolutionary processes might have influenced our ancestors' moral beliefs. I don't intend to give a perfect account of how this works, only to explain enough to get straight to the disagreement between Street and Copp. The following is agreed upon by both of them.

Evaluative attitudes[1] are attitudes of approval/disapproval, and attitudes about what one ought to do (Street, 2006, p. 110). Because these attitudes, and the beliefs or judgements that correspond to them, influence our behaviour, it follows that *certain* evaluative attitudes made those humans that possessed them more likely to survive and pass on their genes to children who survive. A proto-hominid who thinks he ought to kill his children would be very unlikely to pass on his genes. This means evaluative attitudes are potentially just as sensitive to natural selection as physical characteristics like having opposable thumbs; this is a key finding of modern evolutionary psychology (Street, 2006, p. 113).

Street presents the Adaptive Link Account (ALA) of how evolutionary processes (like natural selection) shaped over time what have become our moral beliefs today (Street, 2006, p. 113 onward). Firstly, having certain (primitive) evaluative attitudes in certain situations helps link those circumstances with the kind of response which best promotes evolutionary fitness (i.e. the likelihood of surviving to pass on genetic material). Approving of reciprocal behaviour, for example, could help humans form cooperative alliances in social settings. Over time, humans with the evolutionarily successful primitive attitudes survived, and developed dispositions to make certain evaluative judgements and hold evaluative beliefs (such as 'I should reciprocate if someone helps me) (Street, 2006, p. 127). This accounts for the formation of early humans' normative beliefs.

---

[1] The distinction between attitudes, beliefs, and judgements is unimportant here, so they are used in similar contexts

Copp adds to this account[2]: over time, 'cultural evolution' also occured. As moral judgements and beliefs were passed through generations, groups/societies with certain kinds of norms (e.g. cooperative norms) were more successful. This meant that groups which normalised certain moral judgements survived, and so did those moral judgements.

As demonstrated, Darwinian evolutionary processes have seemingly had a tremendous influence on our evaluative attitudes/beliefs. Realism about value (henceforth 'realism') is that there are evaluative/normative facts or truths 'that hold independently of all our evaluative attitudes' (Street, 2006, p. 110). Realists typically hold that our common-sense evaluative beliefs are at least roughly correct, for fear of scepticism. This means that realists must explain why the content of the attitude-independent evaluative facts, and the content of our moral attitudes/beliefs, are at least roughly identical *given that* our attitudes/beliefs were shaped by Darwinian processes in ways that *do not necessarily* conform to the evaluative 'truth'.

Street argues that the realist faces a dilemma:

1. On the first horn, (Street, 2006 p. 121) realists deny that there is any causal relationship. It is 'extremely unlikely' (Street, 2006 p. 125) that, of the vast array of possible evaluative attitudes, the ones that we have *just so happen* to correspond to the facts. As long as the realist maintains that our beliefs and the facts are roughly the same, they must accept that this as a vast (and remarkably convenient) coincidence. This undermines the plausibility of realism.

2. On the second horn, (Street, 2006 p. 125) they posit that there is a 'tracking' relationship between the two, namely that that 'natural selection so affected our psychology that our moral beliefs tend to track the moral facts'. Copp makes another useful amendment here – this only needs to be a 'quasi-tracking' relationship: our moral beliefs only need to track the moral facts *well enough* that rational reflection could correct for any distorting influence (2008, p. 194).

   The challenge here is that this seems scientifically implausible. The ALA is considered a good, comprehensive scientific theory of evolution. It is unclear why grasping evaluative truths could be helpful from an evolutionary standpoint. It also seems as though the ALA explains the content of our moral beliefs effectively already.

---

[2] Street accepts this addition.

The realist is stuck between an absurd coincidence on one hand, and scientific implausibility on the other. Neither seems a satisfactory explanation of the relationship between the content of the evaluative/moral facts and that of our beliefs (henceforth, the 'facts-beliefs relationship').

Copp's Society-Centred Theory

Copp argues that his 'Society-Centred Moral Theory' (2008, p. 198-201) sidesteps this problem. Note that Copp's approach focuses on *moral* facts/ beliefs not evaluative ones generally, so, this will be the focus of this essay. Copp's theory is as follows: The moral code (i.e. system of moral 'standards' or imperatives) which best serves the needs of a society is the code that is morally authoritative. By-and-large, propositions are *true* or *factual* if a corresponding standard in the moral code has the same content. This means morality for Society S stands in relation to society S. Discussing this theory in detail is outside the remit of this essay; these are merely the key points.

Copp believes this escapes the Darwinian Dilemma, because it implies that evolution pushed the content of our moral facts and the content of our moral beliefs in the same direction (2008, p. 201-203); little wonder their content ends up being so similar:

1. Broadly speaking, the moral code which best serves the needs of Society S are the ones which best promote group evolutionary fitness; in particular, a cooperative and altruistic moral code. This means the moral facts correspond to a moral code that promotes group evolutionary fitness.
2. Similarly, the kinds of moral attitudes and beliefs which tend to survive the process of evolution are the ones which best promote group evolutionary fitness because humans who formed successful groups tended to survive. In addition, groups with more evolutionarily advantageous moral codes (e.g. more cooperation-inducing moral codes) tended to survive the process of cultural evolution described earlier. This meant that over time, the actual moral beliefs that developed got closer and closer to the moral beliefs that best served the group/society's needs (i.e. the moral facts)

Copp (2008) claims that his approach is on the second horn, because there is a 'quasi-tracking relationship' between the facts and our beliefs; they approximate each other *close enough* that it can explain why our moral beliefs today are roughly true. Street (2008) replies that Copp

does not explain how the facts actually cause our beliefs. She essentially accuses Copp of claiming to grasp the second horn of the dilemma, but in reality grasping the first horn, and 'trivially reasserting' (2008, p. 214) the claim that the moral facts have (by complete coincidence) approximately the same content. She claims that he simply assumes 'a large swath of substantive [normative] views' to be the moral facts, 'and then notes that these are the very views evolutionary forces pushed us toward' (2008, p. 214).

This strikes as uncharitable. Copp could reply that his account *does* do something to explain that coincidence; it gives a detailed and nuanced description of how the moral facts are grounded in the needs of a society, needs which in turn are shaped by evolutionary processes. If this account is convincing, there is no wild coincidence that the content of the moral facts and the content of our moral beliefs are similar - this is explained by the fact that both are shaped by evolutionary forces.

Copp claims to be following the second horn and given a convincing explanation of the facts-beliefs relationship; Street claims he has merely restated the first horn and has given no explanation whatsoever. I believe this confusion and disagreement can largely be explained by the ambiguity of the word 'tracking', and the framing of this debate as a *dilemma* rather than a *trilemma*.

Time for a Trilemma

The word 'tracking' (as it is used here) is ambiguous. Street uses it to mean that the facts and our beliefs have the same content, because the former causes the latter (2006, p. 125)**.** In formulating his response, Copp uses 'tracking' to mean that changes in the content of one correlate with changes in the content of the other but not necessarily because the content of the moral facts *causally determines* the content of our moral beliefs (as it would if tracking the moral truths promoted reproductive success, for example) (2008, 194)**.** Copp uses his understanding of tracking (which does not require the facts causally determining our beliefs, but only correlation) in what he calls the 'tracking thesis' (p. 194). Copp recognises that Street takes 'tracking' to imply one causing the other; he describes this as Street's 'tracking account'.

Street accepts Copp's differentiation of the 'tracking thesis' and 'tracking account' (Street 2008, p. 210). However, Street (2008, p. 214) acts as though Copp has accepted the tracking

thesis (i.e. accepted that the moral facts and our moral beliefs are correlated) and rejected the tracking account (i.e. that one causes the other) *without offering any alternative account* of the relationship which could underpin the tracking thesis. However, he clearly does, as I will demonstrate presently.

I will use 'tracking' to mean 'correlation' - simply that the facts and beliefs have the same content. Quasi-tracking, then, means having *close enough* to the same content, or *strong enough* correlation. If we assume that the moral realist needs a quasi-tracking relationship (which is necessary to avoid the sceptical conclusion that most or all of our normative beliefs are 'off-track' (cite Street)) then the question is what causal mechanism underlies this correlation[3]?

The first causal mechanism (on the first horn of the forthcoming trilemma) is 'no causal mechanism'; they are correlated by pure coincidence, much like the infamous correlation between annual swimming pool deaths and the number of films featuring Nicholas Cage. This is what Street referred to in her first horn. The mechanism for the second horn is 'one causally determines the other' either directly or indirectly, much like how the number of pages of Heidegger I read causally determines my will to live (this is a negative relationship). In particular, the content of the moral facts causally determines the content of our moral beliefs. Reverse causality (moral beliefs determine the facts) would not be realism because it would imply that the moral facts are not mind-independent). The mechanism for the third horn is that the content of both is causally determined by a 'confounding factor'. This is a term from statistics which describes a third factor which lurks outside the initial picture, and causally influences both other variables. For example, there is no causal relationship between 'the amount of people who get sunburnt in Dublin' and 'the amount of ice creams purchased from Westmoreland Street Spar'; instead, both are causally influenced by the 'confounding factor' of 'how sunny the weather is'.

Diagrams of all three mechanisms are seen below. The arrows A, B, and C each mean 'causally determine(s)':

---

[3] I assume that we are discussing *natural* moral realist theories, like Copp's. It is notoriously difficult to imagine how irreducibly non-natural facts could have any kind of causal commerce with natural facts. For a thorough (and, I suspect, conclusive) argument against non-natural moral theories, see Street (2006)

1st Horn: Coincidental Correlation

Content of the moral facts

Content of our moral beliefs
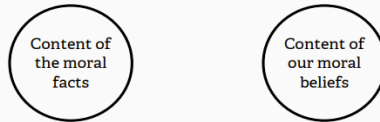
2nd Horn: Causal Relationship

Content of the moral facts → (A) Content of our moral beliefs
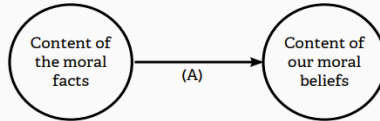
3rd Horn: Confounding Factor

Confounding factor

(B)

(C)

Content of the moral facts

Content of our moral beliefs
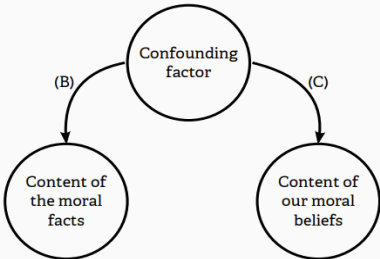
In particular, Copp's approach trades on a 'confounding factor' causal mechanism. His argument is *not* that the content of the moral facts causally determines the content of our moral beliefs, but that *both* the content of the moral facts and the content of our moral beliefs are brought about in a similar way (the ones which promote the survival and success of a society/group survive) and by a similar cause (the environment which societal groups have historically found themselves in). Street seemingly has not considered the possibility that our post-evolutionary beliefs are correlated with the moral facts not because the latter has causally determined the former, but because *both* are causally determined in a similar way by a third, confounding factor. Clarifying this is my first point in this essay: the disagreement between Copp and Street is that Copp thinks it plausible to posit that the causal relationship that determines the content of our moral beliefs so closely mirrors the relationship that determines the content of the moral facts, whereas Street (or someone taking Street's side in the debate) argues that this is implausible.

My second point is to take Street's side in this disagreement, and argue that this 'third way' of explaining the facts-beliefs relationship is not an escape for the natural moral realist. Instead, I argue that it is simply the third horn of a Darwinian *trilemma*. The reason this third explanation is untenable is because it simply pushes the coincidence back a step. The coincidence is no longer simply that the content of our moral beliefs approximates the content of the moral facts (as on the first horn) but instead that the causal process 'B' mirrors the causal process 'C'. In other words, the causal process which explains how the confounding factor determines the content of the moral facts *almost perfectly mirrors* the causal process which explains how the confounding factor determines the content of our moral beliefs. This is, much like before, an astounding coincidence, and a remarkably convenient one for Copp.

Copp's definition of the moral facts is set up such that their content is (in part) causally determined by evolutionary necessity. But it seems as though there are potentially limitless potential definitions of the moral facts, each of which could be causally determined by all kinds of factors. It seems a truly remarkable coincidence (not to mention astoundingly convenient for the moral realist) that the true natural facts just so happen to be the ones which are causally determined in such a way that mirrors the very way in which our moral beliefs are determined.

Ultimately, Street is right that Copp has a massive coincidence to explain (2008, p. 214). My point is that Street misunderstood where that coincidence is located. Street claimed that the coincidence was, just as on the first horn, that the moral beliefs and our moral facts have the same content. Instead, I contend that the coincidence is that the causal process wherein evolutionary forces determine the facts is so similar to the causal process wherein evolutionary forces determine our beliefs. This is a small point. But, in conceding that Copp does *not* merely trivially restate the first horn of the dilemma, but instead offers a causal account of why there is correlation between the facts and our beliefs, the Trilemma is more charitable to Copp. And because this enables it to explain more accurately *why* his approach fails (not because his approach is secretly on the first horn, but because there is another coincidence to explain between the two causal mechanisms he suggests), the Trilemma makes the case against Copp more convincingly. This conclusion generalises to other natural realist theories too; Street's argument was that neither 'coincidental causation' nor a 'causal relationship' are satisfactory ways for realists to explain the facts-beliefs relationship; this essay has clarified that positing a 'confounding factor' is not a satisfactory option either.

Bibliography

Copp, D. (2008) Darwinian Skepticism about Moral Realism. *Philosophical Issues, 18*, 186-206.

Street, S. (2006) A Darwinian Dilemma for Realist Theories of Value. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, 127*(1), 109-166.

Street, S. (2008) Reply to Copp: Naturalism, Normativity, and the Varieties of Realism Worth Worrying About. *Philosophical Issues, 18*, 207-228.

Objects as Sensory Experience: Berkeley's Contribution to the Discussion on Virtual Reality

By Tadhg Cowhig

It may be unthinkable that an 18th century philosopher could have anything to say about virtual reality, but Berkeley's unique philosophical perspective has much to contribute to the topic. Berkeley defends a consistent account of Idealism, which argues that there is no mind-independent substance underpinning the world that we experience. As such, in relation to whether virtual reality is equally as real as concrete reality, his philosophy provides a valid model for virtual reality as real, which maintains our common-sense intuitions. I am not suggesting that the Berkeleyan account works flawlessly in the modern context, but that it helps enlighten us to potential modern philosophy arguments. I will begin this essay by explaining exactly what is meant by the term 'virtual reality'. I will then explain how Berkeley's system can be used to give an account of virtual reality which is equivalent to concrete reality. I will explore an objection to this account, that if we accept its application to virtual reality, then we must accept it for concrete reality as well. Finally, I will show that this objection fails, because we have true beliefs about virtual reality's Berkeleyan aspects which we do not have about concrete reality.

I use the term 'concrete reality', to refer to the world that we currently find ourselves in and normally call 'reality'. I use a specialised term to avoid confusion and prevent biassing our argument by presuming primacy of our current reality. When I describe something as virtual, I mean it in its modern sense of being computer-generated. If I say a virtual okapi, I mean an okapi generated by a computer. (Chalmers, 2017, p.311) Virtual reality is a computer-generated environment which is immersive and interactive. By immersive, I mean that it provides a perceptual experience, so that I see it as immediate in the same way that I experience objects which I encounter in concrete reality. By interactive, I mean that my actions can alter and affect my environment. If I take an action, even if only through an input device like a controller, there should be a reaction in my environment. (Chalmers, 2017, p.313).

Present examples of virtual reality usually take the forms of headsets, such as the

Oculus Rift or the Vive. So far, they have predominantly been used for gaming, but theoretically any form of activity may be conducted through virtual reality. Modern virtual reality is limited compared to the virtual reality that I discuss in this essay. With modern virtual reality, I may believe that an okapi appears before me and it may produce perfectly accurate okapi sounds. However, it will not have the okapi's fragrant aroma and if I try to touch it, I will pass straight through. In this case, it is easy to argue that virtual reality is inferior to concrete reality, because it lacks its sensory information. The virtual reality I discuss, however, requires a future iteration of the technology, one that can provide all the sensory information that we would receive from a concrete object. Whether this level of technology would ever be possible and what form it would take is the subject of a different essay. However, science fiction, from the holo-decks of Star Trek to The Matrix, would indicate that we have intuitions about such a virtual reality, and that we view it as the logical development of current virtual reality technology.

It has commonly been argued in philosophy that even this form of virtual reality is in some sense a lesser form of reality, so that, given a choice between living in virtual reality and in concrete reality, we should always choose to live in concrete reality (Chalmers, 2017, p.309). This aligns with our common-sense intuitions. Most science-fiction stories portray these virtual realities as in some way lesser, with characters having to return to the concrete world. However, our intuitions around future technology are also nebulous, because we have no knowledge about its exact nature. Therefore, we should not take these intuitions to be absolute. I believe that Berkeley's philosophy charts a course for us to construct an argument that virtual reality holds the same value as concrete reality so that there is no reason to choose concrete reality over virtual reality.

Berkeley's doctrine is often summarised as 'esse est percipi' or to be is to be perceived (Downing, 2021, p.2). Essentially this means that there is nothing that can exist in the world separate from our perception of it. Berkeley typically argues for this conclusion, in his dialogues, Philonous taking his position and Hylas that of his interlocutor.

> Philonous: Again, try in your thoughts, Hylas, if you can conceive a vehement sensation to be without pain, or pleasure.

Hylas: I cannot.

Philonous: Or can you frame to yourself an idea of sensible pain or pleasure in general, abstracted from every particular idea of heat, cold, tastes, smells? &c.

Hylas: I do not find that I can.

Philonous: Doth it not therefore follow, that sensible pain is nothing distinct from those sensations or ideas, in an intense degree?

Hylas: It is undeniable.' (Berkeley, 1996, p.113)

Berkeley asks his interlocutor to name any sensible experience, such as heat or sound, and challenges them to abstract from it to any independent object beyond the sensations. He argues that this is impossible because we cannot have pain separate from a particular instance of it. To deny our sensible experience is to deny our common-sense idea of us really perceiving an object through our sensations of it (Berkeley, 1996, p.120). When we cannot, he concludes that there is nothing distinct from our sensible perception of the object because we see nothing 'beside colour, figure, and extension'. (Berkeley, 1996, p.133) Take for example, heat, we cannot imagine heat separate from our sensation of it, something feeling hot. Heat, therefore, only exists inside us as a sensation for a thinking substance.

Of course, in concrete reality, we could argue that heat is just the vibration of particles, but this is not the case with virtual reality. If I was burned by a virtual kettle, though there might need to be a concrete process required to accomplish this effect, nothing within the virtual realm will act to cause our discomfort and we are left with only the sensation itself. Therefore, Berkeley's characterisation appears to capture an important element of virtual reality, the objects that we encounter have no existence separate from the sensations we encounter from them. With the virtual kettle, there is the burning sensation on my hand, the extension and colour we can perceive, the smell, the taste and the distinctive sound it makes but nothing more independent of my mind because we do not expect anything more from a virtual object. Therefore, Berkeley's description of objects appears to align with how we encounter objects in virtual reality.

Berkeley's explanation of mediate perception also aligns with how our intuitions change when interacting in virtual reality. Berkeley argues that:

'For instance, when I hear a coach drive along the streets, immediately I perceive only the sound; but from the experience I have had that such a sound is connected with a coach, I am said to hear the coach. It is nevertheless evident, that in truth and strictness, nothing can be *heard* but *sound*' (Berkeley, 1996, p.144)

This illustrates that for Berkeley we can only perceive objects immediately because the coach is mediated through the walls separating us from it and therefore, for Berkeley we can only strictly say that we have perceived the sound of the coach, which is evident to the senses. Not the coach itself, which can only be inferred from experience, which does not provide us any certainty. Perhaps, we misheard a wagon.

In concrete reality, we can simply refute Berkeley's claim, because we instantly hear the sound and associate it with the coach. Otherwise, we would not be hearing a coach's sound but simply, a strange sound. No further thought beyond the initial sensation of the sound is required. (Dicker, 1982, p.54) However, in virtual reality we are perfectly justified to say that all we hear is the sound of the coach, because there may be no coach producing the sound, it could be a glitch in the software or the coach does not exist at the time the sound is made. Therefore, again Berkeley's philosophy allows us to talk about virtual reality in a way closer to our common-sense intuitions.

The final key element of Berkeley's philosophy that directly impacts our interaction with virtual reality is his conception of God. The common criticism of Idealism is that we have a strong intuition that objects continue to exist in some form independent of me. To circumvent this difficulty, Berkeley characterises his God so that the world is independent of my mind but dependent on the mind of God, declaring 'I on the other side, immediately and necessarily conclude the being of God, because all sensible things must be perceived by him' (Berkeley, 1996, p.152). Therefore, when we are not perceiving an object, it persists because it is being perceived by God. We receive sensations because God wills us to receive sensations. Therefore, it does not matter if we do not perceive objects because they persist in the mind of God and they will remain and be present for us when we require them. When we enter the quad there will be a tree there; it does not matter what happens to the tree while I am outside the quad.

Berkeley's characterisation of the world as sitting in the mind of God seems hard to comprehend in our concrete reality, when science has far better theories for describing our world. However, in the virtual world, it would appear that the role of God is played rather well by the computer. Even when we are not interacting with the virtual world, the computer has it stored in its memory bank, and presents it to us when we need to experience it. The world is quite literally contained within the mind of the computer. Therefore, yet again Berkeley's philosophy maps well onto how virtual reality appears intuitively to us. Therefore, it seems apparent that the philosophy of Berkeley works as in a structurally similar fashion to virtual reality.

The description I have given of Berkeley's philosophy is not meant to be complete or to analyse all its complexities. It was also not meant to establish all possible interconnections between virtual reality and Berkeley's philosophy. I only wished to illustrate that on several principal points, Berkeley's philosophy represents our naïve intuitions about virtual reality and therefore, at the very least works as a good analogy for it. Though Berkeley's philosophy is radical, he was far from the only idealist in his period (Berkeley, 1996, p.IX). However, what distinguishes Berkeley is that his system 'is strong and flexible enough to counter most objections' (Downing, 2021, intro.). Berkeley's position could explain how we interact with the world as well as any philosophy of his day. His system is so strong precisely because he can explain why there is nothing experientially lost in his system, compared to the concrete mind-independent world. If we have all the same experiences, then I see no reason why virtual reality should have any less value than concrete reality. Our encounter with the kettle is the same but the world of matter has been replaced with the world of ideas. For Berkeley's position to maintain its strength, the picture of reality it creates must have the same value as concrete reality. Because we can use Berkeley's philosophy in a modified sense to talk about virtual reality, it would follow that it can have the same value. The virtual kettle has all the same sensory experiences and interactions as the concrete kettle, the only difference is that the former is created by a computer. If we were unaware we were in virtual reality we would not be able to tell the difference between virtual reality and concrete reality, neither appearing lesser. Therefore, Berkeley's philosophy allows us to argue for the parity of value for virtual reality with concrete reality.

However, a sceptic of Berkeley's ability to inform our views of virtual reality might

argue that if we accept the truth of Berkeley philosophy in virtual reality, then we will be forced to accept it in concrete reality. Berkeley believed that his argument applied to concrete reality. If virtual reality has the same value as concrete reality, then any argument that we could apply to the virtual world should be translatable to the concrete world. If we accept the base premises of his argument in the virtual world then we can translate them into the concrete world and maintain their validity, translating references to the computer to references to God for example. Therefore, we are left with the choice either to admit that virtual reality is lesser than concrete reality so we can apply Berkeley's arguments, reject Berkeley's arguments completely, or accept that Berkeley's arguments apply to concrete reality, which no person in good sense would want to do.

I believe that the whole premise of this objection is flawed, virtual reality is in many ways fundamentally different to concrete reality, and therefore not all aspects can be translated from one form of reality to the other. In all examples, when I discussed Berkeley's argument applying to virtual reality, I explained why we had common sense intuitions to doubt these applications in concrete reality, which do not apply in virtual reality. Virtual reality does not need a science of atoms because it is generated out of computer data and therefore, our knowledge of atoms have no relevance to the virtual objects generated.

However, even if the sceptic believes all my arguments around Berkeley's arguments not applying to the concrete world are false, there is still a significant difference between virtual reality and concrete reality, we know that virtual reality is created by a computer. Therefore, we know that we are thinking about a world that is virtual and fundamentally different to the concrete world. In our concrete world, we have no such certainty that a God exists and several arguments to doubt it. Of course God may exist but that does not mean he will have the traits of Berkeley's God, but we can be all but certain that the computer does. Of course, our current reality may not in fact be concrete and instead be virtual, but we are unaware of that fact. Then perhaps, we would have a reason to bring Berkeley's arguments into our concrete world. However, we have no positive reason to believe this, only the inability to completely rule it out. Therefore, I believe that Berkeley's argument can remain solely within the realm of virtual reality.

To conclude, I believe that an 18th century philosopher has a significant amount to add onto our modern discourse on virtual reality. I explained exactly what we should consider

virtual reality and explain why we must refer to some future form of the technology. I then went through several key concepts in Berkeley's philosophy, and explained exactly how they can apply to virtual reality. Immediate objects are reduced purely to the sensory experience removing their mind-independent nature as in virtual reality. Mediated objects are eliminated, preventing us from inferring to them because we cannot infer in virtual reality. As God is the mind which the world is dependent on, so is the computer in virtual reality. I then explained that the strength of Berkeley's argument derives from his ability to offer a description of the world that removes most of our counter-intuitive intuitions about the nature of that world being lesser, which can be used to construct a picture of virtual reality that has equal value to concrete reality. I then presented an objection, that adopting this would require us to adopt a completely Berkeleyan outlook or abandon our project completely. I answered this objection by pointing out that we have certainty, due to knowing the world is constructed by a computer, for adopting such a position in virtual reality that we do not have in concrete reality. I do not believe that what I have written here is in any way a comprehensive account of the ways in which Berkeley's philosophy can inform our research into virtual reality, but merely a brief sketch to establish the validity of the idea. At the very least I believe it shows that virtual reality is perhaps not as new of an idea as we may think.

Bibliography

Berkeley, G. (1996) *Principles of Human Knowledge and Three Dialogue*. Oxford: Oxford
World's Classics.

Chalmers, D. J. (2017) 'The Virtual and the Real'. *Disputatio*. 46, P.309-352.

Dicker, G. (1982) 'The Concept of Immediate Perception in Berkeley's Immaterialism' in
*Berkeley Critical and Interpretive Essays. Ed. C Turbayne. Minneapolis: University of
Minnesota Press. P.83-94.*

Downing, L. (2021) "George Berkeley", The Stanford Encyclopedia of Philosophy (Fall
Edition), Edward N. Zalta (ed.), URL =
<https://plato.stanford.edu/archives/fall2021/entries/berkeley/>.

Gorham, G. (2014) 'Early American Immaterialism: Samuel Johnson's Emendations of
Berkeley'. *Transactions of the Charles S. Peirce Society*. 54, P.441-456. Jones, T.
(2021) *George Berkeley: A Philosophical Life*. Princeton: Princeton University Press.

Kroustallis, B. (2004) 'Berkeley and the Moon Illusion'. *History of Philosophy Quarterly*. 21,
P.151-166.

Nozick, R. (1975) *Anarchy, State and Utopia*. Oxford: Blackwell.

Park, D. (1982) 'On Taking Ideas Seriously' in *Berkeley Critical and Interpretive Essays. Ed.
C Turbayne. Minneapolis: University of Minnesota Press. P.69-82.*

Silcox, M. (2019) 'The Transition into Virtual Reality'. *Disputatio*. 55, P.437-451.

West, P. (2019) 'Reid and Berkeley On Scepticism, Representationalism, And Ideas'. *The
Journal of Scottish Philosophy*. 17.3, P.191-210.

The Possibility of Comprehensible Concepts With Empty Sets: Comparing the Standard Contemporary View with that of Strawson's View

By Jack Palmer

My aunt's biological children are siblings is a true sentence based on the possibility of my aunt having children. If I have never mentioned having any cousins before, you may ask if I do, but until then, you cannot know if my sentence was true or false, yet it makes sense nonetheless. In this essay, I will compare the standard contemporary view on existential import in the four traditional forms of propositions to Strawson's view. In doing so, I will argue that while Strawson is correct about the presupposition of terms, it is possible for more sentences to arise than Strawson allows, dealing with complex concepts and further arguing against the standard contemporary view as Keynes argues for it.

Before we begin this discussion, we must first clarify the four traditional forms of propositions, where they come from, what the issue of existential import is, and then the two key views we will be considering. The four forms make the corners in the traditional square of opposition. The top left of the square is where we see "every A is B" (Parsons, 2008, p. 3). In Aristotelian notation, this would then be noted as "Aab", and in modern predicate logic, this is "$\forall x[Ax \rightarrow Bx]$". The top right of the square is "No A is B", which is represented as either "Eab" or "$\forall x[Ax \rightarrow \sim Bx]$". On the bottom left, we see the claim "some A is B", which can be noted as "Iab" or "$\exists x[Ax \ \& \ Bx]$". Finally, then we have on the bottom right, "some A is not B," which is represented as either "Oab" or "$\exists x[Ax \ \& \sim Bx]$".

The central issue then of existential import arises from an analysis of this square. Traditional Aristotelian logicians would seek to establish relations between each point of the square of contradiction such that we can infer the truth of, say, "Iab" from "Eab" as one example (Parsons, 2008, p. 4). In such a case, we can recognise the two claims to contradict since if some "a" is "b", then it cannot be the case that no "a" is "b" since at least one "a" must also be "b" for us to say some "a" is "b". The critical issue within the square of opposition, though, is that it allows us to say that if we know Eab, then we can take Oab to be true since if

no "a" is "b", then some "a" is not "b" inherently since any existing "a" would necessarily not be "b" if no "a" is "b".

The critical issue here is that we can only infer this if there is an "a" which does exist. This is known as the issue of existential import, being which forms we should interpret as entailing the set of the terms within the proposition to have at least one member which we can take to exist (Parsons, 2008, p.4). This issue, while very well highlighted by the square of opposition, the defenders of which form one of the key positions within the literature, is ultimately an issue which affects any system of logic and what rules of inference we should allow.

The first key view we will be considering is that which is called the standard contemporary view. Proponents of this position argue that universal claims do not entail any existential import, whereas particular claims do entail existential import (Keynes, 1906, pp. 215-218). This is to say, if we were to claim "all A is B" or indeed "no A is B", we are not required to say that there are, in fact, any "A" or "B" and indeed the sets which contain everything which belongs to the predicate "A" or "B" are potentially both empty. This may seem bizarre at first, but we can recognise that we may want to make claims about concepts with no extension, that is, no members within the set of things to which the concept applies. For example, we may want to posit something about things that are not identical to themselves. Separately and far more trivially, though, if I were to say no state is legitimate, we can recognise two ways such a claim could be true. Either if there is a state, then it is not legitimate, or indeed it could just be the case that we say no state truly exists.

However, the standard contemporary view does not seek to do away with existential import entirely. As mentioned, particular or 'existential' claims confer existence. That is to say, if I posit that some states are legitimate, or indeed some states are not legitimate, that necessarily entails that some states exist for them to possess the positive characteristic of legitimacy or the negative characteristic of a lack thereof. One important note here is that if we take this to be the case, then we can no longer infer "some A is B" from "all A is B", as it may be possible that there are no objects which we can take to be "A" or "B", yet if we said "some A is B", that we would be committed to the existence of A. The solution is not to allow the inference from "Aab" to "Iab".

The second view, which we will be contrasting against the standard contemporary view, is that all four traditional propositions imply that the terms in the proposition do not refer to any concept, and the corresponding set thereof is empty. This view is proposed by Strawson, using the example of the statement "all of John's children are asleep" (Strawson, 2011, p. 173). This example is a universal claim, so under the standard view, it does not imply that John has any children. However, Strawson points out that if John had no children, the claim would never arise and, indeed, would be meaningless if it did arise. Hence the form presupposes some existence (Strawson, 2011, pp. 173-176).

The view focuses on applying logic to natural language, which Strawson points out by claiming that the standard view has minimal application, which he deems only formal and not actual real-world applications (Strawson, 2011, p. 170). However, as noted, there are instances where we argue that a concept which we take to exist generally, e.g. that of a state, could be argued to have not genuinely existed, which would have at least clear philosophical applications. Though, if we were to agree that no state existed, Strawson would then claim that any comment about the nature of states would not arise or be senseless.

At this juncture, it is essential to note that the point of contention between these two views is whether or not universal claims result in existential import since both Strawson and the standard contemporary view agree that particular claims confer existence. On this specific matter, I will now turn for a moment to the field of linguistics to begin this essay's main contention. It is important to note where our natural language and even capacity to express arguments in logical notation ultimately come from. Just as babies are not born with innate fluency and command of any language, man was at first equally far from fluency. To this end, I would posit as a primary hypothesis that any concept comes in some way from experience and thus has some object to which that concept is applied.

Within linguistics, the idea of innate language and universal principles across every language is crucial and famously featured in the work of Noam Chomsky. One such fundamental feature in language is the necessity of the noun, the ability to take something to be a subject or object linguistically (Pinker, 2003, p.1-3). In this alone, we see, at the very least, the possibility of every concept referring to an object. Moreover, our capabilities for spatial cognition are seen as crucial in language development, which expresses objects and places in

ways other species do not (Landau & Jackendoff, 1993, pp. 217-218). Within this ability, though, what is especially noteworthy is how natural human language is structured around whole objects and even contains names for object parts (pp. 219). We take the concept of "chair" from our cognition of chairs and how other objects interact with them, but we can also then take the concept of "legs" as they relate to chairs from observing a part of the whole. In each instance, though, the concept we generate is ultimately based on our spatial cognition as it functions according to spaces and objects.

Evolutionary linguistics also then emphasises how syntax structures emerged from the planning of movement using the aforementioned abilities for spatial cognition of both space and objects (Fujita & Fujita, 2022). Natural language necessarily entails concepts relating to objects as linguistics holds that concepts were built to describe certain features of either singular objects or even shared features across a group of objects both now and in the emergence of language. It is important to note, though, that the standard contemporary view is often motivated by the idea of moving past natural language (Keynes, 1906, pp. 226-7). However, the aforementioned linguistic literature emphasises how the features of natural language I discuss are products of the features of our comprehension of experience. This is crucial as we ought to ensure we are not merely playing the game of objective logic while truly comprehending the claims we generate in ways contradictory to the claims themselves.

I also want to point out the faint and somewhat twisted similarity of my claims here to the Fregean concept of extension. One of Frege's fundamental claims and indeed the basic assumption of naive set theory is the idea that any concept has a corresponding set of objects for which the nature of the concept (its 'intension') is the condition to be a member of said set (Burge, 1984, pp. 3-5). The way in which the similarity is twisted is the claim that none of these sets can be empty, and even conceptually, we at least posit an object which then takes on a basic form of existing as soon as we posit a concept.

In this way, we see even the inclusion of even fictional objects. The fact that we can posit Sherlock Holmes as the subject or object of a sentence to which some concept is applied is arguably something which, if logic completely neglects, forces any logical framework down a narrow conception of existence. I can say Sherlock Holmes is a detective and should be able

to represent that without the need for a hamfisted translation into a logical notation using the universal quantifier.

Indeed, I should be able to make claims about John's infamous children from Strawson's example. I should be able to say that John's children are under eighteen. Some may say this is de facto false despite its necessity because John's children do not exist, yet when I say John's children, I do not imagine some ethereal concept. I picture the fictional objects of John's children and recognise them to be under the age of eighteen, just as I take my aunt's biological children to be siblings.

However, these complex concepts combine two or more different concepts or relate to formal concepts, which for Keynes motivates the necessity of the empty set (Keynes, 1906, pp. 240-245). That language allows me to combine the concept of John and the concept of children despite the material non-existence of John's children, which is why we must have an empty set to express such concepts with no extension. Moreover, and arguably more importantly, the concept of that which is not identical to itself is one which human comprehension, imagination, and language have all enabled me to type, which necessarily has a corresponding empty set.

It must be admitted that the latter example is a compelling case against Strawson, as the set of that which is not identical with itself seems to be quite a crucial philosophical concept to investigate. The very non-existent itself has often been featured as a crucial concept within philosophy since Parmenides' poem. However, the claim that Keynes makes of necessarily empty sets is one which I believe is in error.

I want to start this controversial matter with the fundamental conception that even for the non-existent, my expression of the concept involves my linguistic classification of it as a definite subject. To me, this is the first cause to ponder the nature of even imagined yet impossible objects as crucial to constructing both language and comprehension. I cannot even discuss that which is not identical to itself without using the word "that."

As we try to construct a concept whose corresponding set is an empty set, we necessarily construct an imaginary object which would be a member of the set to comprehend and utter the very concept itself. Frege was right to note that every concept has both intension and extension, but was in error to neglect the necessity of extension as something which cannot be empty. In

this way, both the standard contemporary view and Strawson are ultimately misguided. However, Strawson's basic intuition and conclusion are crucial. Not only is it the case that the four forms imply that the terms within each proposition are not empty, but our very ability to formulate any such proposition requires that the terms are not empty.

This is not to say that we can infer that unicorns exist in the material world from the claim that "all unicorns have horns." The very nature of my contention emphasises different kinds of existence, physical or imagined. In this sense, while we should allow for existential import, we must recognise it as implying different kinds of existence based on the context of the nature of the object(s) being discussed.

Concluding on this matter, for logicians who seek to do away with such aspects of natural language and reject my claims, I caution against their hubris in attempting to overcome the nature of human cognition. Secondly, one form of existence could be that of possible existence, possibly motivating modal logic as an olive branch to avoid depending on context. At any rate, though, any proposition necessarily implies that the terms within it are not empty, including that of John's children. On the issue of existential import, then, we must side with Strawson, though in a heavily amended manner.

Bibliography

Burge, T., (1984). Frege on extensions of concepts, from 1884 to 1903. *The Philosophical Review*, *93*(1), pp.3-34.

Fujita, H. and Fujita, K., (2022). Human language evolution: a view from theoretical linguistics on how syntax and the lexicon first came into being. *Primates*, *63*(5), pp.403-415.

Keynes, J.N., (1906). *Studies and exercises in formal logic: Including a generalisation of logical processes in their application to complex inferences*. Macmillan & Company, limited.

Landau, B. and Jackendoff, R., (1993). "What" and "where" in spatial language and spatial cognition. *Behavioural and brain sciences*, *16*(2), pp.217-238.

Parsons, T., (2008). Things that are right with the traditional square of opposition. Logica universalis, 2(1), pp.3-11.

Pinker, S., (2003). *The language instinct: How the mind creates language*. Penguin UK.

Strawson, P.F., (2011). *Introduction to logical theory*. Routledge.

A Way into the World: Heidegger and the Ready-to-hand

By Carla Barry

In this paper, I will argue that in *Being and Time*, Martin Heidegger priorities a form of praxis over theory. I will substantiate my argument by showing that Heidegger thinks that the ready-to-hand is the more prevalent mode of encounter than the present-at-hand, revealing the ontological structure of the world as a totality of references. I will begin by arguing that Heidegger needs to discuss the relationship between the present-at-hand and the ready-to-hand in order to arrive at the phenomenon of the world. In this regard, Heidegger's prioritisation of praxis over theory is clear from his demonstration that Dasein operates with a prior notion of Being-in-the-world which enables us to act in a world that is constituted by a web of references and significations which are not accessible through the present-at-hand.

At the outset, I will briefly define some terms which will be crucial to understanding Heidegger's complex phenomenological account. In this regard, it will be helpful to note that Heidegger introduces his own lexicon, in the interest of conducting a thorough investigation into the nature of Being, whose purpose is to successfully avoid the trappings of the philosophers who preceded him (and who he takes to have obfuscated not only the answer to the question of Being, but the moreover, the question itself).

Perhaps most importantly, Heidegger employs the term *Dasein* (in German, 'Being-there'); this refers to the human way of Being. Dasein is embedded within the world rather than treated as a subject, abstracted from its activities within the world. This re-orientates the structure of enquiry away from the subject-object distinction. To be clear, the subject-object distinction creates a problem of knowledge, as it proposes a fundamental gap between the subject and the world it finds itself within. Contra this distinction, Heidegger's ontological analysis employs a spiral structure. He begins by proposing a return to the fundamental question; what is Being? Here, he argues that Dasein, the privileged entity, functions with a pre-theoretical understanding of Being. An understanding of Being is always accompanied by an awareness that Dasein is fundamentally grounded in the understanding that it must live in the world. The theoretical conception of world as object, as distinct from Dasein, overlooks the primordial connection between existence and Dasein's understanding of existence as Being-in-the-world. Dasein is the privileged entity insofar as it is both ontic

and ontological. Ontic refers to what a thing is. In Dasein's case, Heidegger claims that Dasein's ontical orientation towards the world is concerned with the way it lives, the way Dasein brings itself towards future possibilities of existence. Dasein is the only ontological entity for Heidegger, because through Being, it comes to question the meaning of Being as such. Thus, Dasein is the only entity which can engage philosophically through its movement within the world. In this way, Heidegger prioritises a form of praxis, as Dasein moves through the world in anticipation of future possibilities. The theoretical apprehension of the object, or the present-at-hand, does not account for this. Heidegger's praxis is concerned with the structure of possibility. I primarily encounter objects as objects *of potential use.* In this mode of encounterability, future potentialities for Being come to the fore of my relatedness to the object in question. Through the ready-to-hand, the web of connections between objects in the world, and Dasein's use of them, the totality of references is illuminated. As the ready-to-hand incorporates the understanding that objects always appear within a specific context, equipment is thus embedded with a reference to the context within which it appears. On the other hand, a theoretical understanding treats objects as abstracted form their use, and thus as outside of the totality of references. In the ready-to-hand, the pen refers to the context which is beyond it. In taking hold of an object of equipment, I can grasp that it refers to possibilities beyond the present, and beyond the object's merely physical properties.

At this juncture, the reader may wonder, why is it that Heidegger is interested in the relationship between the present-at-hand and the ready-to-hand? To intimate an answer to this question, we must firstly understand precisely what is meant by these terms. In this regard, we may understand both the present-at-hand and ready-to-hand as modes of Dasein's encounterability with objects. To encounter an object as present-at-hand is to look at it as something which is made up of properties; to describe an object in terms of its being present-at-hand consists of a *theoretical* depiction of what the object is. In this context, I would try to examine the pen as what it is made of, and what form its substantiality takes. The ready-to-hand encounters objects as things which are useful to us, in their *practical* quotidian use. Thus, a ready-to-hand examination of the pen is guided by how I employ it in my everyday absorption in the world.

In turn, this raises the question of why Heidegger is concerned with these two ways of encountering objects. Here, let us recall the question with which Heidegger is primarily concerned: that is, what is the nature of Being? He develops the argument that Dasein, who

exists with an understanding of existence, can only understand Being as Being-in-the-world. This Being-in-the-world for Dasein is not to be conceived of as a spatial insidedness. I cannot say that I am in the world like the bed which is located in my room. In this regard, Stephen Mulhall claims that in traditional ontology, "inhabitation seems like a contingent or secondary fact about human existence, rather than something which is of its essence" (Mulhall, 2005, p. 40). Dasein's Being-in is not to be understood as occupying a space within something. This understanding separates Dasein and world as two distinct entities, where Dasein has just stepped foot into the world. Dasein's Being-in-the-world must be understood as a "'Being alongside' the world in the sense of being absorbed in the world" (Heidegger, 1962, p. 80). Thus, Dasein is fundamentally concerned with the objects it encounters in such a world. Concern is an existentiale, as it is part of the structure of Dasein's understanding of Being. For Heidegger, concern "will designate the Being of a possible way of Being-in-the-world" (Heidegger, 1962, p. 82). Given that Dasein is determined to encounter the world in its concernful dealings, it is difficult to step back and analyse precisely what the world is. Heidegger's aim in this deeper examination of the compound 'in-the-world' is to decipher what the phenomenon of the world is (or the the 'worldhood' of the world). How is it that we encounter the world with a previous understanding that we are already in the world we are trying to explain? This is why Heidegger looks at the relationship between the present-at-hand and the ready-to-hand. This evidences his prioritisation of a form of praxis over theory, since he thinks that the ready-to-hand illuminates the worldhood of the world and its ontological structure. The way that Dasein moves within its everyday environment is primarily praxis. Heidegger uses this practical knowledge to gain a deeper understanding of Being-in-the-world as it pertains to Dasein.

To further my argument, I will look closely at the mode of encounter which Heidegger characterises as presence-at-hand. To see an object as being present-at-hand, the observer tries to intellectually discern what it is. In trying to determine what the pen is in this context, I would explain it as a 'thing' which is made up of a certain substance, displays certain properties and occupies space as an extended substance. In viewing something as present-at-hand I do not use the pen or refer to its function in my actions in the world. In this regard, Magda King argues that the present-at-hand mode "objectivises [things] into mere material substances to be found somewhere in an indifferent universal space" (King, 2001, p. 67). The object then becomes a Thing that lies outside of a world context. Heidegger argues that this "fixed staring" is only one mode of encountering objects (Heidegger, 1962, p. 88).

The object that is perceived as being present-at-hand "can be expressed in propositions, and can be retained and preserved as what has thus been asserted" (Heidegger, 1962, p. 89). He argues that the present-at-hand mode has been used as the primary way of understanding the being of objects which make up the world, in the philosophical method. In turn, this has led to understanding the world as also being present-at-hand, like an object whose properties and form we can perceive. In Heidegger's view, this approach leads to objects and the world being ontologically categorised in terms of "substantiality, materiality, extendedness, side-by-side-ness" (Heidegger, 1962, p. 89). However, these depictions of the being of objects and the world do not adequately show the nature of Dasein's Being-in-the-world. This view of Being-in-the-world depicts "the nature of human contact with the world from the viewpoint of a detached observer of that world, rather than as an actor within it" (Mulhall, 2005, p. 39). This prompts Heidegger to examine the mode of the ready-to-hand. While the object which stands before me must be present and there to encounter, this does not reveal anything about why I am concerned with the object, and the context of the world in which the object exists.

The mode of encounter, readiness-to-hand, is developed by Heidegger since this is the mode which illuminates the phenomena of the world. Dasein's Being-in-the-world is essentially that of concern. Thus, Heidegger examines objects in terms of Dasein's employment of objects as tools which facilitate Being-in-the-world. Dasein's concern is manifest in its various activities in the world; "having to do with something, producing something, attending to something and looking after it, making use of something, giving something up and letting it go, undertaking, accomplishing, evincing" (Heidegger, 1962, p. 83). Dasein's preoccupation and active involvement shows that we do not encounter objects theoretically or present-at-hand. Heidegger uses this claim to argue that objects should be examined through this active involvement as opposed to "a bare perceptual cognition" (Heidegger, 1962, p. 95). When an object is encountered as ready-to-hand, it is used by Dasein to manipulate its environment according to its goal. As Mulhall claims, objects are "employed to alter the human environment; [their] properties of weight and strength subserve the final product, the goal of the endeavour" (Mulhall, 2005, p. 39). The objects which are used in this way are qualified as equipment. Heidegger claims that "equipment is essentially 'something in-order-to'. This in-order-to structure means that the object in use belongs within a specific context of usability. I use the pen in order to write a letter, so that my friend can read it. This way of encountering the pen is such that it is determined by how Dasein uses it

for its purposes in the world. However, in order to understand something as being ready-to-hand for my use, there must already be a larger totality of equipment. This means that I know there is a totality of webs which makes my seeing something as useful possible; "Out of this the 'arrangement' emerges, and it is in this that any 'individual' item of equipment shows itself. Before it does so, a totality of equipment has already been discovered" (Heidegger, 1962, p. 98).

Heidegger claims that praxis is not simply a blind acting. It entails its own way of seeing which is deemed circumspection. In this sense, in my encounter with an object, I also look around the object in order to decipher its place in its environment and its relation to other items of equipment in its locality. Circumspection means to look at the context in which the object operates. The pen appears next to a piece of paper, and is on the desk in the classroom. I can readily identify the relationship between these objects due to my prior knowledge of the contextual structure within which they appear. It is clear from this discussion of the ready-to-hand that Heidegger is coming closer to answering what it is that allows us to experience the world meaningfully, rather than as an unarticulated collection of materials and substances extended in space.

The ready-to-hand maintains itself through being implicit. In our everyday usage of objects, the equipmental totality which makes structured contexts possible, is, for the most part, hidden. I do not stop to question the nature of equipment in my everyday use. However, Heidegger argues that when the ready-to-hand becomes un-ready-to-hand, i.e., not readily available for potential use, the equipmental totality is illuminated. The unusability of an object implies that Dasein operates with a prior awareness of the equipmental totality, as such a totality allows Dasein to encounter objects as objects *of potential use.*

In this regard, Heidegger outlines three ways in which the structure of the ready-to-hand is revealed when equipment is un-ready-to-hand. The first mode of deficiency is when the ready-to-hand becomes conspicuous. This pertains to equipment being broken and unusable. However, Heidegger notes that "we discover its unusability, however, not by looking at it and establishing its properties, but rather by the circumspection of the dealings in which we use it" (Heidegger, 1962, p. 102). The broken pen, which was once a tool, has become somewhat present-at-hand in its state of unusability; "it shows itself as an equipmental Thing which looks so and so, and which, in its readiness-to-hand as looking that

way, constantly been present-at-hand too" (Heidegger, 1962, p. 102). The second mode of un-readiness-to-hand is obtrusiveness. This mode is manifest in instances where some item of equipment needed to fulfil the task at hand is missing. The missing pen brings to light the fact that I need the pen to take down notes in the lecture. The other items of equipment, such as the notebook, "[reveal themselves] as something present-at-hand and no more, which cannot be budged without the thing that is missing" (Heidegger, 1962, p. 102). The final mode of un-readiness-to-hand is obstinacy. Some thing stands in the way and thus interrupts the task at hand. According to Heidegger, such an interruption "enables us to see the obstinacy of that which we must concern ourselves" (Heidegger, 1962, p. 102).Therefore, when a tool becomes an obstacle, I do not want to use it, but rather cast it aside in frustration. These three modes of un-readiness-to-hand show that the tools, when there use-value is compromised, are also material objects which have the "characteristic of presence-at-hand in what is ready-to-hand" (Heidegger, 1962, p. 102). In this regard, however, I think it is important to note that Heidegger does not think that the tool becomes completely stripped of its readiness-to-hand. The broken pen is still acknowledged as a pen as it always belongs to its context of writing, drawing, communication, etc. This is why he argues "that the presence-at-hand which makes itself known is still bound up in the readiness-to-hand of equipment" (Heidegger, 1962, p. 102). Therefore, the present-at-hand is always bound up in the complex web of values and functions in the equipmental totality of the ready-to-hand. Heidegger argues that the present-at-hand is not entirely possible as Dasein cannot step out of the world context which makes objects encounterable. The purely present-at-hand is a type of myth, as Dasein cannot reach a point of pure apprehension and decontextualisation.

Heidegger prioritises a form of praxis through which the worldhood of the world is illuminated. I do not think that Heidegger claims that we operate solely on the basis of practical knowledge. This practical knowledge reveals the ontological structure with which he is interested, when establishing the phenomena of the world. The world cannot just be the totality of objects which appear within it; worldhood of the world makes objects appear as meaningful. Through the interruption of the ready-to-hand structure, the equipmental totality is lit up. This equipmental totality allows Dasein to take up objects and use them when Dasein is concerned with future possibilities for its Being-in-the-world. Therefore, Heidegger thinks that "the context of equipment is lit up, not as something never seen before, but as totality constantly sighted beforehand in circumspection. With this totality, however, the world announces itself" (Heidegger, 1962, p. 102).

For Heidegger, the ontological structure of the world is not accessible through the present-at-hand mode of encounterability. The phenomena of the world are constituted by the sign and referenced structure. This structure entails that individual objects refer to a context to which they belong. The sign and reference structure is necessary for Dasein to interpret the world as meaningful and to operate in the world with active concern for its potentialities; "The ontological structure of world, the worldishness of world as such, is *significance*" (King, 2001, p. 61). Thus, in my use of the pen and notebook, I am not exclusively focused on the task of writing things down. I also think of the possibilities which I strive to fulfil through the present task at hand. I think of the possibility of studying my lecture notes, or writing a future essay. The ready-to-hand mode of encounter is characterised by a "for-the-sake-of-which" (Heidegger, 1962, p. 116). The object I use always points out beyond itself to a totality of equipment which allows Dasein to move through the world and fulfil its possibilities. It is only through this *practical* activity in the world that worldhood as such is made explicit. Therefore, Heidegger claims that the "sign is something ontically ready-to-hand, which functions both as this definite equipment and as something indicative of the ontological structure of readiness-to-hand, of referential totalities, and of worldhood" (Heidegger, 1962, p. 114). On this basis, I think it is clear that Heidegger prioritises praxis over the theoretical understanding of objects. In particular, since praxis illuminates the ontological structure of worldhood, which is constituted by a totality of references, and such a totality makes possible both the ready-to-hand and present-at-hand mode of encounter.

To conclude, I have argued that Heidegger prioritises a form of praxis over theory, as he is fundamentally concerned with what makes the worldhood of the world accessible for Dasein. In examining presence-at-hand and readiness-to-hand, I have argued that the present-at-hand subserves the ready-to-hand, since Dasein necessarily moves through the world with an understanding of context. The equipmental totality enables the emergence of a variety of contexts, which allows Dasein to encounter objects as meaningful, and as tools which can be employed to fulfil possibilities for Being-in-the-world. In this way, I have demonstrated Heidegger's view of praxis, as the approach by which we can understand why Dasein experiences the world as a web of meanings and not simply as a collection of occurrent objects.

Bibliography

Heidegger, M. (1962) *Being and Time*. Translated by J. Macquarrie and E. Robinson. Oxford:
    Blackwell Publishing.

King, M. (2001) *A Guide to Heidegger's Being and Time*. Edited by J. Llewelyn. New York,
    NY: State University of New York Press.

Mulhall, S. (2005) *Routledge Philosophy Guidebook to Heidegger and Being and Time*.
    London: Routledge.

Are There Such Things As Moral Facts?

By Iain MacLeod

In this essay, I will argue that there are no such things as moral facts. First, I will define some key terms. Second, I will outline a criticism raised by Alexander Miller against G. E. Moore's moral non-naturalism: that it reduces to an unjustified intuitionism. Third, I will explain how John McDowell attempts to avoid critiques of intuitionism by using the concept of *Bildung* (i.e., the process of holistic self-cultivation and maturation). Then, I will explain Jerry Fodor's objection, that McDowell does not sufficiently explain Bildung. At this stage, I will present one possible response that a proponent of McDowell's non-naturalism might give; that the process of learning in mathematics as a metaphor for Bildung presents a problem for opponents of McDowell's view. In this regard, I will respond by arguing that axioms cannot be proven in the moral space. To conclude, then, I will consider and reject one objection based on moral consistency, and will end with a brief summary of the ground that has been covered.

A moral fact is a moral statement which is true, discoverable, and mind-independent. For example, if "murder is wrong" is always true regardless of who says or believes it, and we can discover this truth, then that is a moral fact. Moral realism is the view that moral facts do exist, and moral antirealism, the stance which I will defend, is the idea that moral facts do not exist. Moral realism can be divided into two categories; moral naturalism and moral non-naturalism. Moral naturalism is the view that moral facts are reducible to natural facts (facts about the natural world), and moral non-naturalism is the view that moral facts are irreducible. Moral non-naturalism usually aligns itself with strong cognitivism, the view that moral facts are not just truth evaluable, but the way we evaluate their truth is to look at the world outside of ourselves. In this essay, I will be focusing on moral non-naturalism, beginning with Miller's interpretation of Moore's moral non-naturalism:

"For Moore, 'good' denotes a simple, unanalysable, non-natural property, which is not part of the causal order." (Miller 2003, p.35)

Here Miller explains that Moore posits that the good is not a part of the natural world, is causally inert, and is irreducible in the sense that it cannot be broken down into different composite ideas or be shown to be identical to some other property, natural or otherwise. This and most other forms of non-naturalism reject the claim that moral language can be analysed in natural terms; in turn, this leads to the question of how we can analyse moral terms at all. Miller, noticing this, asks more specifically how, under Moore's definition, we can determine which moral concepts are good and which are not. He notes that supporters of Moore's view describe this process as 'intuition' (Dancy, 1991). Intuition, Miller explains, unless it means "the capacity for making correct moral judgements" (which would be explanatorily vacuous), must be some kind of sense-perception which allows us to grasp or detect facts which are not detectable by the senses (Miller 2003, p.36). However, this form of intuitionism, the view that moral facts can be known by intuition, does not adequately explain how we use intuition to ascertain moral facts. More explanatory work needs to be done to show how moral facts can be known, and hand-waving at some special faculty of intuitionism would be, Miller holds, tantamount to supernaturalism. Having outlined these problems with intuitionism, I will now explain how McDowell's moral non-naturalism attempts to combat this problem.

In attempting to resolve this problem, McDowell distinguishes between the space of reasons and the realm of law. The space of reasons is the collection of all meaningful things. The realm of law is the collection of things in natural science which function according to natural laws. In other words, the realm of law is the empirical, observable world. Moral facts, McDowell claims, are present in the space of reasons, and not in the realm of law. To avoid the previous problem of accusations of supernaturalism, McDowell carves a space outside the realm of law that contains (moral) facts which are accessible to us through the process of *Bildung*. This process, he explains, is entirely natural and so is not susceptible to the same charges of supernaturalism that intuitionism is. He writes as follows:

"The threat to truth is from the thought that there is not enough substance to our conception of reasons for ethical stances [...] Earning truth is a matter of supplying something that really does what is merely pretended by the bogus epistemology of intuitionism." (McDowell 1998, p.162)

Here, McDowell proposes that in his conception of moral non-naturalism, we really do have a non-supernaturalistic method of "earning" truth about ethical statements (moral

facts). In so doing, he avoids this charge of supernaturalism which was directed at intuitionism. However, intuitionism also has an epistemological problem: how can we use intuition to access moral facts? Even if these moral facts are in a non-supernatural space, the problem of how we can ascertain truths about them remains. This is where the process of Bildung should explain how we can earn, learn, or otherwise ascertain moral facts. However, as Fodor argues, the description of this process is lacking.

In response to McDowell's claims that practical wisdom "opens our eyes" to moral facts and that ordinary upbringing brings the demands of moral facts "into view" (McDowell 1994, pp.79, 83), Fodor writes the following:

"'Bringing into view' is a metaphor; only what is in Nature [on the usual characterization] can literally be viewed. And 'resonating' is also a metaphor; only what is in Nature [on the usual characterization] can be literally attuned to." (Fodor 1995, p.11)

He argues that McDowell has not given us precise enough descriptions of how the process of Bildung works. While one can imagine that there might conceivably be a way to engage in acts of knowledge-acquisition to ascertain moral facts, similar to how we acquire knowledge of facts in the natural world, using analogous language gives a vague idea of what this might look like but provides neither a proper argument nor a sufficient reason to believe that such a process is actually possible. So, more work needs to be done to support McDowell's view.

Perhaps this work can be done by elaborating on exactly what McDowell means by "earning" ethical knowledge. Miller writes the following:

"For McDowell's cognitivist, too, intuitionism merely helps itself to an unearned notion of truth. But we also have the option of attempting to earn the right to use the notion of truth with respect to moral discourse by providing an account of the sorts of rational criticism to which our sensibilities may be subject, *and in constructing which we may draw on as many distinctively ethical concepts as we find necessary: the right to employ the notion of truth in ethics can he* [sic] *earned by 'working from within'*." (Miller 2003, p.264)

Here Miller describes McDowell's view as being again in contrast with intuitionism, attempting to complete the task intuitionism seemingly could not; to explain our cognitive access to moral facts. Once again, this is attempted via appeal to the process of earning, which involves using ethical concepts which are at our disposal to ascertain truth, and doing so "from within." But what does "from within" mean in this context? Miller writes that "McDowell's use of the notion of cognitive access, like his use of the notion of moral truth, is earned, though from a viewpoint firmly located within a scheme of moral concepts" (Miller 2003, p.264). So, at least on Miller's interpretation, "from within" appears to mean from within the realm or viewpoint of all of the moral concepts that we have at our disposal. Still, reasoning within the realm of moral concepts itself does not seem to be a fully-fledged explanation of the process of ascertaining moral facts. In an effort to make this sound more appealing, I will now consider the idea of using mathematics as a metaphor.

Previously, I discussed McDowell's conception of moral facts as being outside of the empirical, observable world (the realm of law), but nonetheless as being assessable as facts. Mathematics, one might argue, is a (perhaps more accepted) example of something outside of the empirical world, but we can nonetheless assess and ascertain facts about it. For example, a tesseract is basically a 4-dimensional cube; a concept which makes perfect mathematical sense but does not exist in the natural, empirical world. So, perhaps the aforementioned space that McDowell carves out for moral facts possesses similar qualities to this "mathematical space" in which objects such as the tesseract exist. Here, we can make use of metaphor, in the same spirit as McDowell, but in a way which is hopefully more illuminating. When he writes that we should draw from as many distinctive ethical concepts as we find necessary, perhaps this can be understood as similar to how different mathematical concepts may be used to ascertain mathematical truth, some of which are even incoherent or contradictory (like that $1 + 1 = 3$) just as, at least intuitively, some ethical concepts are. In mathematics, we deliberate, without necessarily leaving the mathematical space, about what truths or facts of mathematics there are. This is similar to McDowell's concept of working "from within". Similar methods as those used within the mathematical space could conceivably be utilised to systematically "earn" moral truth through rationalising about the many aforementioned distinctively ethical concepts as might be necessary. So, this is one method that might still avoid some of the concerning problems of how moral truth might

be ascertained whilst still being at least in principle a functional methodology. Bildung, then, could be a similar kind of learning process to the one we use to learn about mathematical facts.

That Bildung has some coherency or familiarity to it in that it possesses these similarities to the process of mathematics is not proof that the process of Bildung can ascertain moral facts in the same way that mathematics can, but it does raise a problematic question for McDowell's opponents; in what way is mathematical learning different from Bildung, such that mathematical learning can ascertain facts, but Bildung cannot ascertain moral facts? In this regard, I will present one possible response to this question; that in mathematics, you can derive axioms.

Every complicated mathematical formula, truth, or theorem is derived from our most basic principles: axioms. For example, for any two numbers a and b; $a + b = b + a$. These axioms are self-evidently true in that asserting its negation (in our case; $a + b \neq b + a$) would be a self-contradiction or otherwise incoherent. There is of course some debate over how axioms are derived, but most agree that there are truths of mathematics, which themselves rely on the truth of axioms (which is all that must be assumed for the purposes of this response[1]). The problem with using the mathematical realm as a metaphor for the moral realm, and then positing that one might be able to use some similar modes of reasoning as are made use of in the mathematical realm in the moral realm, is that it is unclear how one might derive moral axioms. "Murdering for fun is wrong" is a particularly agreeable moral statement and moral concept, but its opposite "murdering for fun is not wrong", while certainly less agreeable, is not self-contradictory or incoherent the way contradictions of axioms in mathematics are. So, because mathematical axioms are derivable and moral axioms are not, the metaphor does not pose a serious problem for opponents of McDowell's view. I will now consider and reject one possible objection to this based on an idea of moral consistency.

---

[1] In the interest of avoiding lengthy discussion about deriving mathematical truth such as is famously present in the hundreds of pages Alfred North Whitehead and Bertrand Russell took to prove that $1 + 1 = 2$, I chose here to appeal to the general consensus that there are mathematical truths to focus more on the subject matter (Whitehead & Russell, 1997).

In this regard, let us consider the complete set of all of the ethical concepts at our disposal – as McDowell suggests that we should, in order to earn, or learn about, moral facts. Some of these concepts will be self-contradictory. For example, "murder is always wrong and it is good to murder Socrates" is a self-contradictory statement. Perhaps the negations of these self-contradictory statements (in the case of the aforementioned murder example, this would be "it is not the case that both murder is always wrong and that it is good to murder Socrates") would constitute moral axioms upon which we can develop objective moral theory just as we have already developed objective mathematical theory, and in this way we can derive moral facts.

However, it is perhaps a mistake to assume that these negations of self-contradictory moral concepts can really be used as a basis for moral facts. Certainly, undertaking the task of ensuring that none of your believed moral statements were self-contradictory would help to ensure that your moral theory is consistent with itself. However, moral consistency does not at all constitute moral truth. Two consistent moral theories could completely contradict each other, but using this method we cannot determine which one is true, only which one is consistent. So, whilst this process may be used to help make moral theories more consistent, self-contradictory moral statements are not axioms in the sense that we can build true theories upon them in the same way we can with mathematical axioms. My conclusion that the mathematical metaphor does not pose a serious problem for opponents of McDowell's view therefore stands, and McDowell's moral non-naturalism fails to show that moral facts exist.

In summation, I began by defining some key terms and outlining Miller's criticism of intuitionism. I then explained how Miller attempts to avoid accusations of supernaturalism by placing moral facts in the realm of reasons and by using Bildung as an explanation for how moral facts can be ascertained. Then, I argued that the metaphor of mathematical learning is similar to McDowell's Bildung in that mathematical facts are outside the natural world, and that mathematical learning uses many concepts within its own realm to reason "from within" to discover or earn truth. I then objected to this process on the grounds that Bildung cannot ascertain moral axioms. I then considered one objection which claims that it can via self-contradictory statements, but rejected this due to this only showing how we can create a morally consistent theory, and not determine moral truth or moral facts. McDowell's non-naturalism is therefore unsuccessful, and so moral anti-realism is true and there are no moral facts.

Bibliography

Miller, A., (2003). *An introduction to contemporary metaethics*.

McDowell, J. (1994). *Mind and World*. Cambridge, MA: Harvard University Press.

Fodor, J. (1995). *Encounters with trees*. London Review of Books 17 (8), 10-11.

Whitehead, A.N. and Russell, B., (1997). *Principia mathematica to* 56 (Vol. 2)*.
Cambridge University Press.

An Exploration of Phenomenological Conceptions of Space and Place Through the Lens of an Advancing Modernity: From Virginia Woolf to the Present

By Isabelle Logan

"Suppose the looking glass smashes, the image disappears, and the romantic figure with the green of forest depths all about it is there no longer, but only that shell of a person which is seen by other people—what an airless, shallow, bald, prominent world it becomes! A world not to be lived in. As we face each other in omnibuses and underground railways we are looking into the mirror that accounts for the vagueness, the gleam of glassiness, in our eyes. And the novelists in future will realise more and more the importance of these reflections, for of course there is not one reflection but an almost infinite number".

– The Mark on the Wall. (Woolf, 1917)

Virginia Woolf is renowned for her insightful and perceptive analyses of the world and how we relate to it. Much of her writing incorporates subtle messages regarding the importance of place, which in turn intimates her views on the ways that landscapes and spaces affect the people within them. In this regard, Woolf pays marked attention to the technological advancements which occurred during the interwar period of her writing, observes modern developments, and ponders the influence that these are likely to have on future generations. Furthermore, in *A Room of One's Own*, Woolf considers in-depth the importance of space, where personal identity is a concern – an issue which I will focus on in greater detail in the discussion that follows.

Woolf's deep appreciation for and understanding of the importance of space, place, and distance implies her sophisticated phenomenological[1] conception of the way in which we identify ourselves in our surroundings. In this essay, I will endeavour to articulate this conception, drawing on the analyses of academics who have studied the phenomenological aspects of her work in detail. I will compare (my interpretation of) Woolf's conception of how we identify ourselves in our surroundings with a variety of contemporary accounts, from Marc

---

[1] Phenomenology is concerned with analysing the phenomena characteristic of direct experience.

Augé, Alberto Sanchez Rojo and Michelle Perrot. Furthermore, I will consider whether Woolf was correct to predict that advanced communications, globalisation, and supermodernity would play vital roles in the process of modern self-identification.

One particularly interesting aspect of contemporary writing on connectivity and modernity and Woolf's writing – both of which deal with phenomenological concepts of space and place – is the consideration that each affords to technological advancements, and how these are liable to influence our perception of ourselves.

As evidenced by the above quote, from *The Mark in the Wall*, Woolf's consideration of technology, and the ways that we can recognise ourselves within it, often subtly permeate her writings. One example that stands out can be found in her novel, *Mrs. Dalloway*, when the author uses the bus (referred to then as the 'omnibus'), as a metaphor for escaping what she terms the "damned egotistical self" (Schroder, 2007). While most commuters and travellers would likely consider a journey on the bus to be a mundane experience, about which there is little to philosophically ponder, Woolf's outlook allowed room for her characters to see an image of the self, as this manifests within the mundanities of daily travel. In this way, Woolf illustrates that an escape from the ego can be conferred when the physical self is hoarded onto a vehicle, which is in the process of transporting two dozen others through space, time, and a bustling city.

In *The Mark on the Wall*, Woolf further considers the humbling experience of the individual within the herd, within the vehicle, when she captures the experience of the fleetingness and fluidity of the narrator's surroundings on the London underground, when they are "being blown through the Tube at fifty miles an hour" (Woolf, 1917). This leaves the reader with the idea that the advent of motor transport marked the first instance in which humans opted to surrender their physical selves – if only for brief periods – to a machine. Interestingly, embodying transport in this way implies something of a reciprocal transaction between *us* and machine, insofar as the machinery of transport requires human direction and interaction just as much as the human requires its service.

This fusion of physical body and machine inevitably significantly impacts a person's self and body-image, and has the power to alter the way that they view themselves, as the machine plunges them into motion, transporting them from one location to another in just a

few short moments. The philosopher Leena Kore Schroder develops this idea in *Reflections in a Motorcar*, and considers the interactive nature of the relationship between us and the world. In this regard, Schroder highlights a monistic approach to the way that our perception of the world cannot be separated from our perception of ourselves, so long as we are actively participating in the world. She highlights Woolf's frequent allusions to busses and motorcars to substantiate this idea, and claims that it is "a realisation that is enabled by the car because its motility actively places the body into a dynamic relation with time and space" (Schroder, 2007, p. 134)

This monistic approach to phenomenology conceives of the self not as an individual body, but as a property of the world around it, surpassing the "mirror stage" of being; this rejects the objectification of the self. In allowing us to explore a plural and communal self, we are encouraged to consider ourselves as being one with the world and think of the self beyond the limits of self-alienation, and the confines of personal identity, which so often delude a person's ego and promote an unhealthy relationship with their surroundings.

While Woolf was among one of the first generations to observe the merging of the self with the machine, few others shared her insightful views on what this would mean for future generations as technology developed. Today, a huge part of the process of self-identification, and the way that we identify those around us is, is intimately linked with our technologies, if not inseparably so. With the development of smartphones and constant access to the internet, we have online spaces and communities through which to express ourselves and communicate with others. Not only do we curate idealised versions of ourselves to construct and maintain an image of a person we wish to be perceived as by others, furthermore, we begin to identify with this image, so much so that a typical social media presence serves as a better personal resume than a physical meeting would.

Instead of dwelling on the complicated nature of digital spaces, I will instead focus on the possibility of intertwining visions of ourselves with our modern surroundings and advancements in technology. This is a concept developed by Marc Augé, via his concept of "non-places". In *Non-Places: Introduction to an Anthropology of Supermodernity,* the anthropologist outlines his theory that some places, or preferably, locations, are so far removed from the experience of being a human in the world that they are instead to be named "non-places". The difference, according to Augé, between places and non-places, is that

non-places do not hold enough significance to be regarded as actual places, in the same way that anthropologically defined places are places (Augé & Howe, 1995). It seems that by the time Augé was writing in 1992 the novelties of modernity that had so appealed to Woolf and her characters had worn off and people were beginning to focus much more on that emptiness than the escapist attachment to such public and man-made spaces.

In non-places, like train stations, hotel rooms, shopping centres and supermarkets, humans are nameless passers-by, anonymous actors who neither relate to nor identify with their surroundings. Somewhat liminal, non-places are spaces constructed by humans with a purpose, for example, an airport, where there is a constant mood of surveillance and security, and a total sanitisation of all residual remains of human interaction with the space. These spaces are generally defined by their lack of character and constant limiting of human interaction, as they appeal to the supermodernist lifestyle, which is characterised by artificial surfaces, security, and an impatience that brings forth the acceleration of time (Anon, 2009).

In a slightly complicated way, the phenomenological insight to takeaway from Augé's theory of non-places is similar to that manifest in Woolf's comments just discussed. Both writers think about the contemporary spatial production of their world in a way that undermines the individual ego, entrenches conformity, and encourages invisibility of the individual. Both conceptualise the rules of movement and behaviour in these public spaces in a way that removes individuality from one's relation to the world.

For Woolf, this comes in the form of her use of these spaces as a way to convey a sort of escapism for her characters from their egos, while at the same time highlighting their inner alienation from the material world around them. For example, in *Mrs Dalloway*, both Elizabeth and Clarissa Dalloway are described as finding comfort and solace from the world while sitting and observing taxis and omnibuses going by in the city; they are, at the same time, pondering the meaninglessness of the very ego that they are escaping. In *Mrs Dalloway*, Clarissa puts a pause on a shopping trip to observe bustling cityscape and is described as having "had a perpetual sense, as she watched the taxi cabs, of being out, out, far out to sea and alone; she always had the feeling that it was very, very dangerous to live even one day" (Woolf, 1925).

While Woolf revels in her description of the brief but liberating detachment from the

egotistical self, moving as one with the temporal flow of surroundings, Augé's view seems to resent the lack of character, restriction of movement, and panoptic security of such non-places. Although it seems that Augé refrains from commenting on the morality of non-places and does not infer any explicit disdain or fondness for them, it appears that – from his language, and the very term in discussion being a negation of the term 'place', (which is anthropologically defined, by him, as valuable in terms of identity, and as a space where people can engage with their surroundings and socially interact) – there may be a lack of enthusiasm on his part for the possibility of phenomenological discoveries within no-places.

As mentioned above, the modern age of information and the concept of supermodernity comes with a certain excitement for the acceleration of time. Almost every technological advancement today aims to save its user either time or money, its putative counterpart. Time and space become tightly controlled, as demonstrated by the idea of non-places, and as described by Alberto Sánchez Rojo in his journal on *The Formative Value of A Room of One's Own and its Use in a Hyperconnected World*, there is "a space for every time, a time for every space and a particular way of behaving in each" (Rojo, 2018).

In this paper, Rojo outlines the effects of our co-dependent relationships with technology through highlighting the importance of actual solitude and isolation from them if we wish to accomplish a sense of self-identity or even self-discovery. He incidentally draws on the topics dealt with in Woolf's *A Room of One's Own*, and highlights the importance of having a room where a person can rest, think, and create within the confines of their own physical space, as well as their own minds, with no outside influences detracting from that experience.

While Woolf's theory of the room of one's own deals with the importance of a room from the perspective of an oppressed woman and her expression of her own identity as defined by her gender, Rojo's article promotes the concept of the room from the perspective of a person seeking isolation from modernity. He considers the positive effects a room with only a desk, a couch or a bed, and a window might have on a person's self-identity in our supermodern world. Here, we can identify an unexpected connection between Woolf and Rojo, between the importance of place and personal space for individual experience.

In this regard, both writers see the room as necessary for its user to be at one with

their most creative and honest self. It is seen as a fundamental and formative space in which its user can thrive. However, one major difference between the two lies in the fact that for Woolf, while the room is a fundamental aspect of what permits the emergence of creativity and writing, one can only assume that she would have been of the opinion that the user would benefit from modern technology such as radios or gramophones. According to Schroder, her diary reveals how she marvels at such devices, as on one occasion when she writes "we opened one little window when we bought the gramophone… And now another opens with the motor" (Schroder, 2007). In a similar vein, in *The Waves*, Woolf describes Louis' experience with modern communications as he claims to be "half in love with the typewriter and the telephone" as he is fusing his "many lives into one" (Lewty, 2007). From this, it is evident, as discussed previously in this essay, that Woolf's admiration for these technologies, however insightful she may be otherwise, perhaps lacks the foresight to see the negative effects that their predecessors have had on individuals' relationships with the world around them.

Today, and as evident in Rojo's paper, the effects of modernity on the individual's sense of self in the world seem to attract quite a bit of attention. While we do often talk about the negative effects of hyperconnectivity and the age of smartphones, it is less often that we consider the effect that they have on our sense of identity. It seems that if having a space to call one's own in some private corner of the world is essential in helping a person to reconnect with their surroundings, and to have an enlightened corporeal experience, it must be best considered without the aid, distraction, or influence of modern technology. That said, it is pertinent to wonder whether the person is then keeping up with their surroundings, or really living "in" the world surrounding them; should they extricate themselves from a major aspect of the essential nature of our hyperconnected, modern world?

In this regard, the French author and historian, Michelle Perrot, considers the importance of space to a person's expression and exploration of individuality in *The Bedroom: An Intimate History*. Through collections of secondary sources and cuttings from novels, she conjures up an image of the importance that a person's personal space has on their psyche (Perrot & Elkin, 2009). While considering socio-economic and cultural differences, Perrot hypothesises that, where possible, it is most advantageous for a person to inhabit a room of their own for reasons of self-discovery, self-actualisation, and a deepened understanding of

their role in and relationship with the world that they live in.

Overall, the theory of the importance of a personal space in which to flourish, as asserted in 1929 by Virginia Woolf, stands strongly today not only as a spatial theory of significance, but also one of phenomenological import. Almost one hundred years since, her reflections on the impact that space can have on an individual's identity and understanding of the self remain vital. Although her enthusiasm for contemporary technology may not be as popular a perspective today as it was then, evidenced by the ostensible rise in disdain many of us have for hyperconnectivity and supermodernist spaces and technologies, there is nevertheless much to be said for having a room of one's own – with or without advanced and up-to-date technologies.

In my view, the concept of non-places serves to remind us of how many of the places we frequent in our daily lives are devoid of distinctive character which bears on our lives and selves. Although they are often necessary to visit, and difficult to avoid, it appears that there is value in adopting Woolf's approach: viewing them appreciatively, and trying to derive value from their fluidity and communal essence.

Bibliography

Anon, (2009). *Archaeologies of Place: Places, non-places and supermodernity: on the issues of rooting and uprooting.* [Online]
Available at:
https://www.brown.edu/Departments/Joukowsky_Institute/courses/archaeologiesofpla ce/7994 .html [Accessed 04 03 2023].

Augé, M. & Howe, J., (1995). *Non-Places: Introduction to an Anthropology of Supermodernity.* Croydon: Verso.

Dunn, L., (2020). My Sister, My Mirror. *Aeon*, 23 10.

Lewty, J., (2007). Virginia Woolf and the Synapses of Radio. In: A. Snaith & M. H. Whitworth, eds. *Locating Woolf: the Politics of Space and Place.* London: Palgrave Macmillan.

Perrot, M. & Elkin, L., (2009). *The Bedroom: An Intimate History.* London: Yale University Press.

Rojo, A. S., (2018). The Formative Value of A Room of One's Own and its Use in a Hyperconnected World. *The Journal of Philosophy of Education,* 53(1), pp. 48-60.

Schroder, L. K., (2007). Reflections in a Motorcar: Virginia Woolf's Phenomenological Relations of Time and Space. In: A. Snaith & M. H. Whitworth, eds. *Locating Woolf: The Politics of Space and Place.* New York: Palgrave Macmillan, pp. 131-148.

Woolf, V., (1917). The Mark on the Wall. In: D. Bradshaw, ed. *The Mark on the Wall and Other Short Fiction.* Oxford: Oxford World Classics, pp. 3-11.

Woolf, V., (1925). *Mrs Dalloway.* London: Alma Classic Ltd.