6th PhiGS Colloquium

5–6 June 2024

University of Kent, W1-SR6 (Woolf College)

Book of Abstracts
Contents

The Conceivability of Consciousness—Kirk's and Chalmers's Zombie Thought Experiments

In this paper we will discuss *Zombie TE* (thought experiments) from (e.g.) Robert Kirk and David Chalmers. On rhetorical analyses there appear three possible fallacies, popular science fallacy, objectivist fallacy and straw man, which are restorable to some extent.

On surface analyses of Kirk's *Zombie Replica* we discover one more TE, *Zulliver*, an alternate of *Brain-in-a-Vat (BIV)*. On deep analysis as by Kirk himself in *Stanford Encyclopedia of Philosophy* it is a *Conceivability Argument*, that could be considered basic to TE in consciousness studies. Because of complexity and modal structure of *Conceivability Argument* we discuss a nonmodal, non-TE example as *inverted spectrum* (ST1), that appears acceptable to skeptics as W.V. Quine. Chalmers proposes his global *Zombie World TE* as an argument from failure of logical supervenience of the mental on the physical. Chalmers's *Anti-materialism Modal Argument (AMMA)* appears on deep analyses with help of modal logic an ontological modus tollens, falsifying materialism.

As nonmodal alternate argument (ST2) we next discuss Saul Kripke's *Pain ≠ (□) C fibers Firing*. It is itself not a TE, but as according to Chalmers most essentially based on related *Divine Creation TE* by Kripke, which Kripke interpretation has received less attention than merited. Lastly, we discuss a non-TE example of *phantom pain* as mentioned by Kripke (a/o) which appears acceptable to Daniel Dennett, who even accepts René Descartes's metaphorical explanation of the *bell-pull*, although Dennett maintains it is only an example of a relatively 'thin hallucination.' We end the paper with some suggestions of further research as based upon Chalmers's catalogue of conscious experiences.

Keywords:
thought experiment, antifallacy, zombie, conceivability argument, anti-materialism modal argument, qualia, pain ≠ C fibers firing, divine creation, modal semantics, substitution thesis, inverted spectrum, phantom pain

Robert Kirk, David Chalmers, Saul Kripke, Roy Sorensen, WVO Quine, Daniel Dennett

05-06 11:20 | **BRIAN KETT** | UNIVERSITY OF KENT

For AI, Models of Understanding and Explanation are Different to Models of Precision

With the increasing use of AI in many fields, there has been a demand for explaining the outcomes it reaches. Thus, the research area of explainable AI (XAI) has been avidly pursued. Despite this, attempts to garner explanations from AI models have had little success. The best results to date have come from surrogate models providing local explanations. Even here, the results have been mixed. One possibility is that surrogate models have been based upon models designed for optimising accuracy, whereas a model to understand or explain an outcome has a different aim—to explain. Virtues of explanation are explored here to use as better criteria for development of suitable explainable models. Even so, newer AI models that use deep neural nets may not be explainable at all and it is argued that explanation should be replaced by trust as the criterion for accepting AI use.

05-06 12:10 | **KIMBERLY MADERO** | UNIVERSITY OF COLORADO BOULDER

Addressing Bias in Facial Recognition: Implications for AI and Future Technologies

Facial recognition technology is widely used in a plethora of industry and government sectors, from law enforcement to consumer applications and health care support. A growing body of evidence suggests these systems exhibit significant racial biases. In this paper, I explore the challenges and consequences of facial recognition inaccuracies when applied to minority populations and the corresponding impact on the development of artificial intelligence (AI) and future technologies.

Biases in facial recognition algorithms are widely documented, with studies revealing categorically higher error rates when the technology is presented with individuals of non-white skin tones. The highest concentration of errors occurs when the algorithm is presented with people of darker skin tones. Because these inaccuracies contribute to the exacerbation of systemic discrimination and unjust profiling, such systems raise clear ethical concerns. In particular, the underrepresentation of diverse input in the datasets used for training the algorithms is identified as a major contributor to the biases in facial recognition systems.

As facial recognition technology continues its permeation into society, its shortcomings pose significant risks to minority populations. For instance, the inability to recognize facial expressions in the healthcare industry among people of color leads to demonstrably negative patient outcomes as the introduction of robots to help autistic children becomes more normalized.

Such issues can compromise the privacy and civil liberties of individuals while simultaneously contributing to increasing racial and ethnic disparities. This paper discusses the implications of facial recognition bias on marginalized communities and calls for a comprehensive reevaluation of the technology's development processes.

Several solutions to mitigate the racist nature of poorly trained algorithms have been proposed, including the diversification of datasets, increased transparency in algorithmic decision-making, and regulatory frameworks that promote fairness and accountability. However, the inexorable integration of facial recognition technology into all areas of society demands that the ethical implications of inherent racial bias in these algorithms be urgently addressed. An urgent need exists for a paradigm shift in the development and deployment of facial recognition technology to prevent it from exasperating the existing societal inequities. B actively working to understand and rectify these biases in every stage of development, we can ensure that AI contributes to the creation of more equitable and inclusive technologies, fostering a future where innovation aligns with ethical considerations to promote respect for the dignity and rights of all individuals regardless of their skin color or ethnic background.

The Relation of Leibniz's Law of Continuity to the Reality of Matter

This essay discusses a version of Leibniz's Law of Continuity which can be seen as a natural law holding that all objects in motion change gradually rather than discontinuously. Leibniz's Law of Continuity is both a natural and a metaphysical law. Leibniz's Law of Continuity relates to his metaphysical conception of the reality of matter because objects in motion are real or actual material objects. The primary outcome is a contribution to the secondary literature on Leibniz's Law of Continuity. I will consider two of Richard T.W. Arthur's claims of Leibniz's notions of continuity and argue that they are mistaken. I contend Arthur's first mistaken claim is Leibniz's conception of continuity as solely an abstract or ideal property. I will argue that Leibniz's view of continuity is best described as a gradual change of an actual object in motion. The heart of my argument is that Arthur's view of continuity as a solely ideal property does not seem to explain the foundations for Leibniz's Law of Continuity since this law applies to actual objects in motion. One reason for this is Leibniz's belief in the existence of the only substances, called "monads", which must exist in all of creation. I hold that Arthur's second mistaken claim is Leibniz initially believing that matter consists of infinitely many ideal points. I will argue that Arthur's second claim is contradicted by two of Leibniz's initial views. Firstly, early in his career Leibniz believes that an actual object in motion must be ultimately composed of infinitely many actual parts of matter. Secondly, at first Leibniz specifically claims that space, which consists of matter, does not contain any points.

05-06 14:50 | **MATTHEW NICHOLAS COATES** | UNIVERSITY OF CALIFORNIA IRVINE

Does it Harm Science to Suppress Dissenting Evidence?

With a rise in spurious (epistemically detrimental) dissent in science, for example climate change denialism, there have been discussions amongst social epistemologists about what should be done by scientific communities. One proposed method to deal with this spurious dissent is by hiding it, such as by hindering it during the peer review process. However, the effectiveness of such a strategy has been questioned, for example by de Melo-Martín and Intemann (2018), in part because we do not have a good characterization of exactly what dissent is spurious.

Using network models, following the paradigm developed by Zollman (2009), I simulate the effect that scientists hiding dissenting evidence without such a characterization has on the beliefs of scientific communities. Computational network models have not previously been used in debates around epistemically detrimental dissent. My model fills this gap in order to offer new possible explanations for observed phenomena and to lend further support to arguments that are already in use in this debate.

I find that it is typically epistemically harmful to hide dissent. Importantly, I find that the epistemic harm caused by hiding dissent holds even when there does exist a biased agent who is purposefully producing spurious evidence to support the dissenting theories. This is important as it is the case most focused on in debates about epistemically detrimental dissent, where it may be expected that hiding dissent would stop the biased agent. This may support the views of de Melo-Martín and Intemann that we should not hide dissent to combat spurious dissent. I show these results are robust under other modelling assumptions, beyond those typical of the results of Zollman (2009), and other modelling frameworks.

However, I do find it may be epistemically beneficial to hide dissent when scientists are already sufficiently far along the path to reaching consensus. In this case hiding dissent is as successful at reaching correct consensus, whilst also doing so more quickly. Quicker consensus may be important in situations where fast decisions must be made, such as public health emergencies. This may justify hiding dissent when the mainstream theory has large amounts of evidence supporting it.

References:
de Melo-Martın, I. and Intemann, K. (2018). *The fight against doubt: how to bridge the gap between scientists and the public.* New York: Oxford University Press
Zollman, K.J.S. (2009). "Optimal Publishing Strategies". Episteme, 6(2), 185–199.

Beyond Market Participation: Revaluating Economic Freedom and Morality in Misesian Economics

Ludwig von Mises, a prominent figure in the Austrian School of Economics, is a staunch advocate for economic freedom, asserting that individual liberty inherently extends into the economic domain. His broader economic theory emphasises the importance of free markets as the most efficient allocators of resources, driven by the self-interested actions of individuals. Central to his philosophy is the idea that voluntary exchanges in a free-market system create a complex network of prices that convey essential information about consumer preferences and resource scarcity, ultimately guiding optimal resource allocation in society. However, Mises's conceptualisation of individual freedom seems to hold a utilitarian rather than an intrinsic value. He argues that freedom is essential because it enables self-interested choices that underpin the establishment and efficient functioning of a free-market society, thereby fostering productivity and overall societal advancement. This essay critiques Mises's interpretation of individual freedom, labelling it as a form of possessive individualism that distorts the true meaning and significance of the notion of individual freedom.

I argue that while Mises champions economic liberty as a cornerstone of a prosperous society, his interpretation of individual freedom is narrowly construed as the capacity to make market-oriented decisions reducing therefore, the rich and multifaceted concept of freedom to mere market participation. This reductionist view potentially overlooks the broader ethical and philosophical dimensions of individual autonomy and moral agency. The analysis considers Mises's most representative writings on economic freedom, such as Liberalism: The Classical Tradition (1927) and Economic Freedom and Interventionism: An Anthology of Articles and Essays (2007). By examining these works, the essay aims to elucidate Mises's perspective on economic freedom and provide a comprehensive critique of his rugged economic individualism.

| **HUGH ROBERTSON-RITCHIE** | UNIVERSITY OF KENT

Explanations for Epistemic Injustice in Healthcare Encounters— Evidence from Chronic Fatigue Syndrome.

I aim to give explanations for epistemic injustice (EI) in healthcare encounters, particularly explanations of how and why health care professionals are persistently perceived to be the agents of epistemic injustice (Kidd & Carel 2016) despite their training and their good intentions. My own background as a family doctor may help me gain insights into the perspectives of other medical professionals.

Philosophers' contributions to this debate concentrate on the analysis and evaluation of patients' experience of and health care professionals' (HPCs') contribution to that epistemic injustice. Philosophers cite "asymmetries of experience and knowledge, and asymmetries of power relations" (Carel and Kidd 2007, p. 336.)

I see an explanatory connection between epistemic uncertainty and epistemic injustice (Blease et al. 2017). Because of their medical knowledge, training, and experience, HPCs are likely to have different conceptualisations from the patient of the pathological processes that are making the patient ill, and which treatments should be recommended. I suggest that epistemic injustice in CFS follows from these epistemic mismatch problems which are prominent in CFS because of uncertain and disputed knowledge of the pathological processes that precipitate and perpetuate its symptoms.

06-06 11:00 | **DAVID MATTHEW** | UNIVERSITY OF KENT

Role Ethics and Consequentialism

In this paper I address a key question for contemporary role ethicists. Can a role-based theory of our ethical lives set itself apart from its major rivals, deontology, consequentialism, and virtue ethics, or will it inevitably collapse into one of these more popular positions? Role ethicists claim that though it may share certain features with other ethical theories their view is nonetheless unique. I believe that they are correct. I will restrict my attention to just one of role ethics major rivals, consequentialism, and will argue that while role ethics is able to accommodate some of the core intuitions that lie behind consequentialist ethical theories, it is also substantially distinct. Therefore, it seems safe to conclude that role ethics cannot be reduced to a form of consequentialism, despite some notable similarities.

A Role for Resource Situations in the Semantics of Generics

Grant, for the sake of argument, that all bare plural generics have a simple subject- predicate Logical Form. Then we have the following analysis:

(1) (a) Dogs bark. $\Rightarrow$
     (b) bark(dogs)

What is the reference of 'dogs' in (1a)? According to orthodox versions of this view, it is the kind dogs (e.g. Carlson, 1977; Liebesman, 2011). Whether or not this is correct, we can ask how it would come to be the case that bare plural nouns in English refer to kinds. One option is that they are simply names of kinds, analogous to singular proper names. An alternative, which I will pursue here, is that bare plurals refer to maximal sums, in the same way as plural definite descriptions. The question discussed here is what is included in that sum.

One of the great innovations of the situation theory of Barwise and Perry (1983) was showing how situations could be used to restrict the reference of singular noun phrases. Following this insight, I adopt a situation semantics for plural definite descriptions, extending Elbourne's semantics for singular definite descriptions (2013), where the definite article has the following semantics, taking a covert situation pro- noun to restrict its reference:

$$\text{lthel} = \lambda f_{<e,st>}.\lambda s: s \in D_s \land \exists x \big(f(x)(s) \land \forall y (f(y)(s) \to y \leq x)\big).\sigma x f(x)(s)$$

'$\sigma x f(x)(s)$' refers to the maximal element of f in a situation s (Link, 1998) and $\leq$ is the part-of relation of Classical Extensional Mereology. I show how variation in the situation pronoun, known as the *resource situation* (Cooper, 1996), can account for some of the puzzling variation in the meaning of generics.

Constraints on resource situations across the following dimensions affect the interpretation of generics:

(i) **Space.** Generics cannot take very small resource situations. For example, 'lions have manes', uttered in front of a cage of lions cannot mean that the lions in that cage have manes.

(ii) **Time.** Generics can take resource situations which are extended in time, into the past and future. This explains why 'dogs bark' is intuitively about what dogs in the past, present and future do; it cannot be true in virtue of only the current crop of dogs barking.

(iii) **Alternatives.** Resource situations for generics are sensitive to the alternatives to a sentence. For example, the resource situation for 'ducks lay eggs' can include only situations of ducks producing offspring in some way.

References:
Barwise, J. and Perry, J. (1983) *Situations and attitudes*. Massachusetts, USA: MIT Press.
Cooper, R. (1996) 'The role of situations in generalized quantifiers', in S. Lappin (ed.) *The handbook of contemporary semantic theory.* Oxford: Basil Blackwell, pp. 65–86.
Carlson, G.N. (1977) *Reference to Kinds in English.* Ph.D Thesis University of Massachusetts Amherst.
Elbourne, P.D. (2013) *Definite descriptions.* Oxford: Oxford University Press.
Liebesman, D. (2011) 'Simple Generics', *Noûs,* 45(3), pp. 409–442.
Link, G. (1998) 'The Logical Analysis of Plurals and Mass Terms: A Lattice-theoretical Approach', in Link, G., *Algebraic semantics in language and philosophy.* Stanford, California: CSLI Publications, pp. 11–34.

06-06 13:40 | **HOWARD LIU** | UNIVERSITY OF MIAMI

Knockdown Arguments in Philosophy, Deep Disagreements, and Philosophical Success

Knockdown Arguments in Philosophy, Deep Disagreements, and Philosophical Success In non-philosophical disciplines like physics, chemistry, or geology, it is not hard to find arguments that reach decisive conclusions. Anyone who objects to premises or supporting relations between premises and conclusions in these arguments is often regarded as being irrational or not able to fully understand the argument. Call these arguments knockdown arguments. As conclusions from knockdown arguments have been established, these disciplinaries can further develop based on these conclusions. What about philosophy? Is there any knockdown argument in philosophy as well? Although philosophers do argue in the tone as if they are proposing knockdown arguments for various theses as well, it seems less straightforward whether there are knockdown arguments in philosophy if we take prevalent disagreements in philosophical communities and constant debates about philosophical success or progress into consideration. While philosophers like Lewis (1983) and van Inwagen (2009) deny the existence of knockdown argument in philosophy, on the other hand, Ballantyne (2014) attempts to argue that we can obtain knockdown argument in philosophy from some established non-philosophical knockdown arguments.

In this paper, I argue for a skeptical position on whether we know the philosophical arguments at hand are knockdown or not due to epistemic underdetermination in philosophy. First, I argue that Lewis and van Inwagen's rejection of knockdown arguments' existence in philosophy is untenable since the position is either self-refuting or not a position that we are rationally compelled to believe in. Second, I present Ballantyne's strategy that attempts to establish knockdown arguments in philosophy from conclusions established by non-philosophical knockdown arguments and Keller's criticism (2015) that arguments generated from Ballantyne's strategy beg the question. I argue that Keller's criticism only scratches the surface of the issue and also has the consequence of failing to retain the conviction that there are knockdown arguments outside philosophy. The deeper problem of Ballantyne's argument, I contend, is that it is always possible for an actual (or potential being) to rationally deny one of the premises that Ballantyne deploys. Here, I appeal to literature about deep disagreements (Ranalli & Lagewaard 2022a 2022b) to support my position. That is, it is always possible to disagree with certain theses or claims in philosophy without seriously being irrational. However, I argue that even if we should endorse a skeptical position in knowing knockdown arguments in philosophy. We need not be pessimistic about the future of philosophy since the notion of philosophical success and knockdown arguments in philosophy are independent of each other.

References

Ballantyne, N. (2014). *Knockdown Arguments*. Erkenntnis, 79(S3), 525-543.

Keller, J. (2015). *On Knockdown Arguments*. Erkenntnis, 80(6), 1205-1215.

Lewis, D. (1983). *Philosophical Papers, Vol. I.* New York: Oxford University Press.

Ranalli, C. and Lagewaard, T. (2022a). *Deep Disagreement (Part 1): Theories of Deep Disagreement.* Philosophy Compass, 17(12), e12886.

Ranalli, C. and Lagewaard, T. (2022b). *Deep Disagreement (Part 2): Epistemology of Deep Disagreement.* Philosophy Compass, 17(12), e12887.

Van Inwagen, P. (2009). *Metaphysics (3rd ed)*. Boulder: Westview Press.

Poor Iggy the IGUS Loses Her Faith in the Future

It is widely agreed that we experience the world from a temporal perspective, but it is widely disputed whether that is because there is a changing temporal order, or merely because we are the kind of creatures that are built to experience things that way.

Callender (2017) claims that manifest time, the regimented common-sense picture, is due to us being IGUSes; Information, Gathering and Utilising Systems. Accordingly, all that is required, is epistemic and psychological explanations rather than metaphysical ones. We can attribute manifest time to the kinds of creatures we are. This works great at describing how we might have acquired manifest time but where do we go from here? While the IGUS account works well at explaining why we are predisposed to manifest time it does not offer practical alternatives to the innate beliefs that support IGUS systems.

I consider the example of pathological temporal experiences due to depression, as described by Ratcliffe (2015), to show that a practical/ motivational relationship to time requires committing to the metaphysical asymmetry between past and future. The elimination of the possibility space of the future coincides with an erosion of the self but given Callender's account, this would be a more accurate/objective relation to time. Callender thinks that 'the self' is an illusion and this illusion is responsible for why we accept manifest time.

What kind of normative demands can these 'facts' of scientific time make? Specific cases of depression have an impoverished U. They experience thoughts but they don't think well. As one person put it "I don't have any real ideas. I am led by "idea-memories." These idea-memories lack conative power. The obvious response is to fully support the depressed realism thesis. However, what these examples suggest is that the loss of 'self' and manifest time results in a cognitive impairment. What if we imagine the entire scientific project as an IGUS system? would there be a corresponding loss of affordance of objects in the environment that drives the scientific project itself?

References:
Callender, C. (2017). *What makes time special? (First edition.)*
Ratcliffe, M. (2015). *Experiences of Depression: A Study in Phenomenology.* Oxford University Press.

06-06 15:40 | **ELLIOT PORTER** | UNIVERSITY OF BRIGHTON

Moral Dark Matter

Around 30% of the observable universe is supposed to be made up of dark matter. We infer that this matter exists because of its gravitational effects, however it seems not to interact with electromagnetism in any observable way. We infer it's existence, though we cannot 'see' it. I suggest that the same is true in the moral sphere – that the world contains a great many reasons that have genuine normative force, but which we cannot 'see' in the usual way we 'see' our practical reasons. These reasons make up a significant body of moral dark matter.

The argument proceeds from two premises. The first, controversial, premise is a strong form of normative externalism. Many of the facts of the world have an irreducible to-be-doneness about them. The second, probably non-controversial, premise is that human minds offer only a limited perspective on the world. We have epistemic access to only some of the facts that are out there. Humanity enjoys only a narrow window on the (normative) world, leaving much out of sight, and a penumbra around of edge of things only-just-visible. From this, I infer that there are some irreducibly normative facts which lie beyond human ken, and probably some that lie just at the periphery, just where our reach exceeds our grasp.

It will be objected, of course, that these facts cannot be reasons if they lie beyond our ken, either because we could not in principle respond to them, or because – being unparsable in human terms - they cannot count as facts in the first place. I respond; first, that such facts would remain irreducibly normative, and so still make up moral dark matter; second, that facts at the periphery could be responded to, however imprecisely, and so remain reasons; and finally, that unparsable facts are fine as long as we distinguish facts qua propositions from facts qua things specified by propositions.

Finally, I consider what kinds of facts might make up this peripheral moral matter, and how we might come to know it. I suggest that facts about the more-than-human world, its value, and how we should relate to it, are prime candidates. I suggest that atypical mental states, particularly manic and mystic states, might count as specialised modes of cognition that offer privileged access to this otherwise opaque moral knowledge.