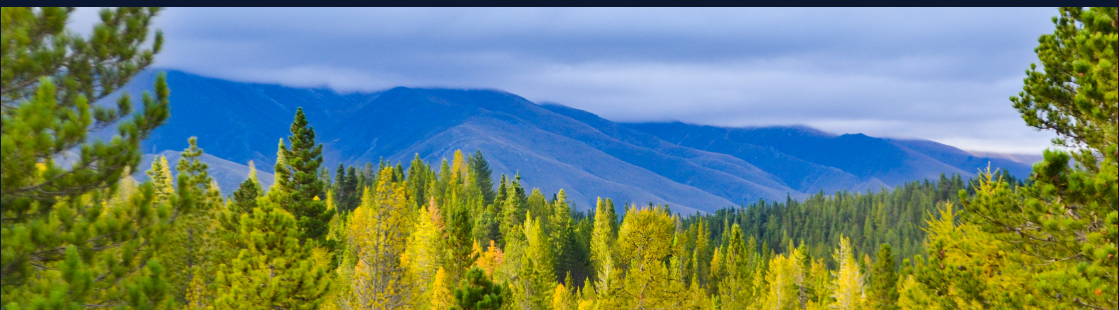


Issue 11 (1) Winter Conference 2017

*British
Journal of
Undergraduate
Philosophy*



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

Journal of the British Undergraduate Philosophy Society

British Journal of Undergraduate Philosophy

The Journal of the British Undergraduate Philosophy Society

Issue 11(1) Winter 2017

Editor-in-chief: Ella Langham

Assistant editors: Bede Hager-Suart
Nathan Oseroff

Manuscript editors: Harriet Ainschough
Tommy Hewlett
Sophie Osiecki
Theodore Stone
Ng Sai Ying
Shike Zhou

Commissioning editor: Edmund Smith

President of BUPS: Eugene Chua

British Journal of Undergraduate Philosophy

www.independent.academia.edu/BJUP

All rights reserved.

© 2017 British Undergraduate Philosophy Society, and the authors.

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

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

Editorial

The greatest weariness comes from work not done.

— Eric Hoffer

I'm incredibly proud to present this edition of the British Journal of Undergraduate Philosophy and furthermore proud to be the journal's first female editor. This edition has been no mean feat. A changeover of the committee has meant that those involved have had to learn quickly and take on roles above and beyond what we were prepared for. Although the task has been daunting at times, it has been a delight to work with the new team and now finally being able to sit back and admire the fruit of labour that is this edition. It has been a pleasure to be working with Eugene Chua, Victor Roulière and Edmund Smith in the months upcoming to the publishing of the journal and for their tireless work and cooperation. We could not have succeeded in this transition without having been supported by the wisdom of veteran committee members: Farbod Akhlaghi-Ghaffarokh, Geoff Keeling, Achilles Saradaris and Rajeev Dass, to whom all of us at the journal wish all the best in their future endeavours.

A massive thank you must also be said to the team of manuscript editors: Tommy Hewlett, Harriet Ainscough Needham, Shike Zhou and Conor Thompson-Clarke who have worked enthusiastically despite the tight time-scales. A special thanks also to Bede Hager-Suart, Nathan Oseroff and Ng Sai Ying for devoting their time to L^AT_EXing. A particular mention must be made of Sophie Osiecki to whom we are immensely grateful as, despite having enough work on her plate tirelessly L^AT_EXing, has also taken on the position of Finance Officer and without whom this conference would not have been possible. Further, we would like to take this opportunity to thank UCL for allowing us the space to host the conference and to the University of Durham for their generous funding. We are also very excited to have Robin Halpin as our keynote speaker at the Winter Conference.

On a personal note, I must extend a further thanks to my dear friend Bella, who has supported me through all the ups and downs of preparing for the publication of this journal.

This edition includes papers ranging from Krike's modal argument to Kant and showcases some of the best undergraduate work. Despite the challenges facing its creation, this edition features some fantastic pieces and we are very excited to have some of these outstanding undergraduate philosophers present their work. I hope you enjoy reading it as much as I have enjoyed working on it.

EL

Contents

iii. Editorial

1. **Explanatory Gap and Psychological Explanation**
Benedetta Romano
12. **Conceiving ‘Ideally’ and its implications for the Modal Argument**
Ahmad Jabbar
20. **Against Atomism**
Gareth Pearce
27. ***Post-hoc* Memory Erasure as a Response to the Kane–Widerker Objection**
Harry Sherwood
34. **Kant’s Epistemology of Testimony**
Anni Raty
48. **Solipsism in the Tractatus**
Sam Davies Udina
56. **A Host of Uncontroversial A Priori Knowable Contingent Truths**
Oliver Hurcum

67. **Has Davidson rejected conceptual relativism?**
Anne Deng

Explanatory Gap and Psychological Explanation

Benedetta Romano

Ludwig Maximilian University of Munich

1 Abstract

According to the explanatory gap argument, psychological phenomena cannot be explained properly, because they cannot be necessarily deduced from neuroscientific events. I argue that the problem posed by the explanatory gap is tied to the definition of explanation typical of the Deductive-Nomological (DN) model, which requires a bottom-up necessitation of mental properties from neurophysiological features. I claim that this assumption is not adequate for a field like psychology and its explanatory practices, as they are better fulfilled by a mechanistic kind of approach. I maintain that the latter can offer satisfactory explanations without being necessarily deductive. In virtue of this, I argue, a mechanistic type of explanation removes the gap, because it removes the requirement that, in order to make intelligible our mental functions, they should be necessarily reduced to the brain's physical characteristics. I conclude that, since the explanatory gap problem depends on the assumption of the DN definition of explanation, by abandoning it in favor of the mechanistic type, psychology can avoid the explanatory gap.

2 Introduction: What is the explanatory gap and how does it concern psychological phenomena?

The term 'explanatory gap' was coined by Levine,¹ referring to the epistemic problem of understanding the relation between mental and physical phenomena. Even admitting, Levine argues, that mental states may indeed be realised in physical ones from an ontological perspective, the epistemic problem of how the distinctive features of mental events can be explained in terms of physical properties remains. Levine's main focus was on the qualitative characters of conscious experiences (viz. qualia). However, the explanatory gap concerns

¹Levine [19].

the explanation of mental functions as well, such as perception, memory, reasoning and emotion, as well as human behaviour.²

Levine claims that facts about conscious states cannot be properly explained by physical facts because they cannot be derived a priori from them in the way that facts about water or temperature can be. To illustrate this difference, Levine makes use of an epistemological version of Kripke's metaphysical argument against materialism,³ comparing physical identities with psychophysical identities. According to Levine, the crucial distinction between regular scientific identity statements, such as 'temperature is kinetic energy', and mind-brain identity statements, such as 'pain is C-fibre firing', lies in their explanatory character: while the former is fully explanatory, the latter is not. Identifying temperature with kinetic energy explains all the properties of temperature since the causal role associated with temperature is played by the kinetic energy and, consequently, the two things are identical. Differently, identifying pain with C-fibre firing explains something about pain, but not all the properties of pain: specifically, the physical process (viz. C-fibre firing) by which the causal role of pain is effected, cannot explain the qualitative character of pain, that is: how pain feels.

Since the concept of pain includes more than its causal roles, the psychophysical identity leaves us with an explanatory gap that is absent in the case of temperature and kinetic energy. The lacuna described by the explanatory gap

²The idea behind the explanatory gap was not born with Levine. Before him, Leibniz argued in *Monadology* §17 that the knowledge of the operations of a physical mechanism could not account for the corresponding subjective perceptions: "It must be confessed moreover that perception and that which depends on it are inexplicable by mechanical causes, that is, by figures and motions. And, supposing that there were a machine so constructed as to think, feel and have perception, we could conceive of it as enlarged and yet preserving the same proportions so that we might enter it as into a mill. And this granted we should only find on visiting it, pieces which push one against another, but never anything by which to explain a perception."

Similarly, Locke, in *An Essay Concerning Human Understanding* Book IV Chapter III, sustained that our sensations and their respective ideas could not possibly be deduced from the physical processes which triggered them; the connection between primary and secondary qualities was not necessary, but arbitrarily decided by God, who, for instance, determined which particular configuration of particles gives rise to a certain perception of colour: "We are so far from knowing what figure, size or motion of parts produce a yellow Colour, a sweet Taste, or a sharp Sound, that we can by no means conceive how any size, figure, or motion of any Particles, can possibly produce in us the Idea of an Colour, Taste, or Sound whatsoever; there is no conceivable connexion betwixt the one and the other. . . . [T]he Ideas of sensible secondary Qualities, which we have in our Minds, can, by us, be no way deduced from Bodily Causes, nor any correspondence or connexion be found between them and those primary Qualities which produce them in us . . ."

³Kripke [18].

originates from the lack of a necessary connection between physical properties and mental phenomena as, for instance, one can conceivably imagine C-fibres firing without pain, and vice versa. In Levine's words:

So long as the nature of that qualitative character is not explained by anything peculiar to any particular physical realisation of pain, we have no way of knowing whether or not a different physical realisation of pain, in a different creature, is associated with the same qualitative character.⁴

What goes for the identification of pain with neurobiological events goes for any identification of conscious mental states with physical events, as described by physical and functional kinds of explanations.⁵

More generally, as Harman⁶ notes, the explanatory gap regarding conscious states is just an instance of the difference between objective and subjective understanding: the first is characteristic of the physical sciences, which aim at understanding phenomena 'from the outside', while the latter concerns the sciences of the mind, broadly conceived, such as psychology and the cognitive sciences, which deal with phenomena that cannot be completely accounted for through an objective approach, and require an understanding 'from inside'.

Numerous and opposing reactions to the explanatory gap argument have been proposed, leading to diverse consequences for the state of physicalism.⁷ The present paper will not discuss any of them in particular, but will focus on a peculiar epistemological aspect of the explanatory gap problem, that is how the latter is crucially determined by the definition of explanation adopted in

⁴Levine [19].

⁵Besides Levine's, many arguments for the explanatory gap have been formulated, aiming to show that phenomenal properties cannot be satisfactorily explained in physical terms. Among the most famous are the arguments by Chalmers [3] on the conceivability of zombies, Jackson [15] on Mary the colour scientist, and Nagel [22] on what it is like to be a bat.

⁶Harman [14].

⁷Responses to the explanatory gap's argument included the claim that it reflects a limitation of our cognitive capacities, as we are constitutionally unable to bridge the gap (e.g. McGinn [23]); others have claimed that the gap shows how the notion of qualitative phenomena is flawed, and qualia do not ultimately exist (e.g. Dennett [10]); alternatively, according to Hardin [13] and Clark [6], alleged explanatory gaps are historically contingent and can be closed as scientific knowledge advances; further strategies included conceiving the gap intuition as resulting from psychological processes: as such, the gap would persist even when confronted with a satisfactory physical explanation (Fiala [11]). Finally, Tye [25] and others have argued that the explanatory gap is expected, given the discrepancy between physical and phenomenal kinds of concepts.

relation to this problem. Since the phenomena involved in the explanatory gap are those studied by disciplines like psychology, the following paragraphs will focus on the explanatory practices characterising this discipline, and will argue that the gap is tied to the reductive kind of explanation typical of the Deductive-Nomological (DN) model. I will conclude that since approaches other than the DN model seem more suitable for providing adequate psychological explanations, psychology can avoid the problem posed by the explanatory gap.

3 How the Explanatory Gap is tied to the DN model of explanation

The explanatory gap argument does not deny the existence of correlations between phenomenological experiences and brain processes. Neither does it deny that progression in the field of neuroscience may disclose them. However, according to the argument, these correlations would only be contingent and neuroscience, therefore, could not reveal anything about the nature of conscious experiences: to this purpose, in fact, not only is it necessary to know that pain is correlated with C-fibre, but also that it must be so, in such a way that the phenomenological characteristics of pain are a priori implicated by the firing of C-fibre. This conclusion stems from the notion of explanation adopted by the argument, which conforms to the Deductive-Nomological model: according to this notion, the explanandum, in order to be explained, must be deducible from the explanans.

More specifically, according to the DN model, which Levine and the other proponents of the explanatory argument endorse, explanation is conceived of as a necessary explanatory reduction: the explanation of a phenomenon is nothing other than the deduction of the phenomenon to be explained from a lower-level explanatory science. In other words, the explanandum is explained when it is necessarily reduced to the explanans, that is, when the laws governing the behavior of the explanandum are deduced from the laws governing the behavior of the explanans. For instance, an adequate explanation of why water boils at 100°C must show how this phenomenon can be deduced from (i.e. reduced to) the laws of chemistry and physics. Since conscious states cannot necessarily be deduced from physical ones, any explanations of the former in terms of the latter are resultantly incomplete.

Hardin and Clark have argued that the contingency between neurological facts and phenomenological experiences is only apparent, and is due to our limited

access of neuroscientific data at this point in time. For Hardin,⁸ there are reasons to be optimistic about the possibility of shrinking the gap in the future. Our neuroscientific knowledge can already explain a considerable number of phenomenological features: studies on colour vision, for instance, seem to suggest the presence of important asymmetries between different colour hues, which have a distinct phenomenological character grounded in the underlying physiology. Along the same line, Clark points out how the existence of explanatory gaps is a historical phenomenon: at the time of Locke, for instance, the liquidity of water, which we can explain now, defied explanation just as much as the relation between primary and secondary qualities.

We think there is an explanatory gap for sensations, but how do we know we are not in Locke's shoes? . . . At the time we lacked an analysis of the notion of water which would mesh with a suitably detailed physical story in such a way as to explain how water solidifies. But, some three centuries later – *voilà!* – the gap is closed. This line should give us grounds for optimism. . . . Human experimental psychology has had barely a century to get underway.⁹

Neither Hardin or Clark seem to question the idea that a reductive kind of explanation can be provided for psychological phenomena, and that this will close the gap. As the explanatory gap argument requires a bottom-up necessitation of mental phenomena from neurophysiological features, the claim that a greater knowledge of neuroscientific facts may close the gap seems defeasible. Cummins,¹⁰ in his examination of the existing explanatory paradigms in psychology and their vulnerability to the explanatory gap, claims that only the so-called strong neuroscience approach to psychological explanation would be immune: according to this approach, as proposed by Churchland,¹¹ only the “concepts generated by neuroscience proper to articulate its data and theory should be used to reconceive the mental from the bottom up, discarding those mentalistic concepts that have no clear neuroscientific reconstruction, and simply replacing those that do”. “Psychology on the mental side of Leibniz's Gap (i.e. the explanatory gap) will either be assimilated or perish.”¹² The strong neuroscience approach would close the gap because it establishes mental

⁸Hardin [13].

⁹Clark [6].

¹⁰Cummins [7].

¹¹Churchland [5].

¹²Cummins [7].

properties to originally be described in pure physical terms, and it therefore assimilates them to their physical features.

However, as Cummins explains, this approach is infeasible, because “the vast majority of explananda (in psychology) are formulated in terms that either explicitly or implicitly draw on concepts that have no known counterparts in neuroscience...”. “Marvelous as the new technology is, it does not, and cannot, provide ‘psychology glasses’, lenses through which observed brain anatomy and activity emerge as psychological faculties and thought processes.”¹³ The explanatory frameworks offered by BDI (belief–desire–intention) and computational approaches to psychology, for instance, explain psychological phenomena through a vocabulary different from neuroscience’s: the gap arises because the bio-neural description of processes in the brain cannot be translated into folk psychology and the computational characterisation of psychological processes.

The lack of correspondence between concepts offered by neuroscience and those offered by the other psychological frameworks echoes the distinction, proposed by Levine between physical (i.e. ‘thin’) concepts, which are based on the description of the causal role played by their referent, and phenomenological (i.e. ‘thick’) concepts, which are based on the immediate quality of experience that presents itself.¹⁴ According to Levine, the incompatibility between thin and thick concepts renders psychophysical identities unintelligible (i.e. unexplainable), even if they were true, because thin concepts cannot be derived from thick ones, and therefore the latter cannot explain the first.¹⁵ If we extend this idea beyond phenomenal experiences, to psychological phenomena in general, the following results: the content of a concept formulated by psychological frameworks other than neuroscience cannot be explained given the respective neuroscientific concept, and therefore the explanatory gap arises. This conclusion, however, reflects the idea that to make the relationship between psychological concepts and neuroscientific concepts intelligible, a necessary deduction of the first in terms of the latter is required. We may wonder ‘Why would one expect that phenomenal concepts/thick concepts should be based on the descriptions of the causal/functional roles in order to be fully explanatory?’¹⁶ As we have seen, this depends on the assumption of the DN

¹³Cummins [7].

¹⁴Levine [20].

¹⁵Ibid.

¹⁶Kostić [17].

model of explanation. In the next paragraph, I will illustrate how this assumption is not the most suitable for psychology and its explanatory practices.

4 Why the DN model does not suit the explanation of psychological phenomena

Aside from the criticism directed to the DN model in general, what is relevant for our discussion is that this framework does not seem to fit a field like psychology. Firstly, psychology is defined as a special and not a fundamental science like physics because it does not deal with nature in general, but with a special system (viz. the mind) with a peculiar constitution and organisation. As such, mental phenomena cannot be explained by general laws of nature, but rather in terms of their specific constituents and the way in which they are organised.¹⁷

Moreover, psychology broadly conceived as the study of the mind is constitutionally a multidisciplinary field as it works with data drawn from different levels of analysis. Hypotheses on psychological phenomena are formulated within distinct paradigms, which are independent of one another although mutually supportive, at least potentially, in considering each others' results to regulate their own procedures.

Finally, it should be considered that psychological explanation, as any other kind of explanation, must provide 'epistemic satisfaction' in order to be successful.¹⁸ Campbell draws on Davidson's idea that the intentionality and the rationality of an action cannot be explained by its physical characteristics: a physical description may present the event as a necessary consequence of its physical cause, but this gives no insight into the intentional or rational character of the event, as only a mental description of the causal players can accomplish this.¹⁹ In the same way, in order to achieve epistemic satisfaction in the explanation of mental phenomena, we cannot turn to their physical description. Otherwise, we would incur an explanatory gap.

Even though mental events are physical events, explanations must be couched either in mental or physical vocabulary (not both). To insist that from a physical description alone we can discern that an

¹⁷Fodor [12].

¹⁸Campbell [2].

¹⁹Davidson [9].

event is the rational outcome of a deliberation is simply to expect too much from physical theory, for this ignores the fact that events must be type identified in ways that allow us to achieve epistemic satisfaction.²⁰

The characterisation of psychology as a special science, its interdisciplinary constitution and its demand for epistemic satisfaction seem to be better taken into account by frameworks other than the DN model, such as the mechanistic type of explanation.²¹ Firstly, differently from the DN model, mechanistic explanations do not necessarily involve general laws or conceptual and logical connections among the elements of the mechanisms.²² Secondly, for the DN model, the relations between distinct fields such as neuroscience and folk psychology are just conceived as reductive, but mechanisms in mechanistic explanation are multi-layered. The structure of a mechanism can be described at a high level, with a specific vocabulary and methodologies, whereas descriptions of a lower level might require different tools of investigation and vocabulary.²³ Finally, the mechanistic framework can offer a satisfactory explanation without being necessarily deductive: it removes the explanatory gap because it removes the requirement that, in order for our mental functions to be made intelligible, they should be deduced necessarily from brain physical characteristics.

5 Conclusion: Psychology can avoid the explanatory gap

Adopting a mechanistic model of explanation dissolves the explanatory gap because this model recognises that neuroscience and other explanatory paradigms refer to the same psychological phenomenon, without imposing the features described within other paradigms to be reduced to those detected by neuroscience. In this respect, the role of neuroscience in relation to psychological explanation is no longer to bridge the explanatory gap by providing a description of psychological phenomena in pure neuroscientific terms. Instead, Cummins argues, Neuroscience may contribute to psychological explanation by functioning as a source of evidence for theories formulated by other explana-

²⁰Campbell [2].

²¹See Bechtel and Wright [1] for a more detailed definition of the mechanistic account, and how it can provide a unifying framework of what is psychological explanation.

²²Taylor [24].

²³Craver [8].

tory paradigms.²⁴

In the same way, the results achieved by paradigms other than neuroscience may guide neuroscientific developments. In this regard, Clark significantly notices that the progress leading to close Locke's explanatory gap on water concerned not only the increase of knowledge in physics, but also the very notion of water as determined by the current state of physics as a field.²⁵ This idea is further elaborated by Chomsky, who compares the past situation between chemistry and physics to the current relation between psychology and neuroscience.²⁶ He recalls the fact that, before the quantum revolution in physics, Russell denied the possibility of chemical laws' being reduced to physical laws. Afterwards, however, says Chomsky

[T]he new physics was unified with a virtually unchanged chemistry, but there was no reduction in the anticipated sense . . . , they were unified by radically changing the underlying science. That could very well be the case with say, psychology and neuroscience. I mean, neuroscience is nowhere near as advanced as physics was a century ago Unification may take place, but that might require radical rethinking of the neurosciences, perhaps guided by computational theories of cognitive processes.²⁷

In conclusion, the characteristics of psychology as a field seem to elude the reductive type of explanation postulated by the DN model. Since the explanatory gap problem depends on the assumption of the DN explanatory framework, by abandoning this framework in favor of a more suitable one, such as the mechanistic type of explanation, psychology can avoid the explanatory gap.

References

- [1] Bechtel, W. & Wright, C.D. (2009). 'What is psychological explanation?'. Retrieved from <http://philpapers.org/archive/BECWIP>.

²⁴Cummins [7].

²⁵Clark [6].

²⁶Chomsky [4].

²⁷Chomsky [4].

- [2] Campbell, N. (2009) 'Why we should lower our expectations about the explanatory gap', in *Theoria*, 75/1: 34–51.
- [3] Chalmers, D.J. (1996) *The conscious mind: in search of a fundamental theory*. Oxford: Oxford University Press.
- [4] Chomsky, N. (2014) 'Science, Mind, and Limits of Understanding'. Available at https://chomsky.info/201401__/.
- [5] Churchland, P.M. (1981) 'Eliminative materialism and propositional attitudes', in *The Journal of Philosophy* 78: 67–90.
- [6] Clark, A. (1994) 'I am Joe's explanatory gap'. Available at <http://selfspace.uconn.edu/paper/PGAP.HTM>.
- [7] Cummins, R. (2000) "'How does it work?'" versus "'What are the laws?": Two conceptions of psychological explanation', in *Explanation and Cognition*: 117–144.
- [8] Craver, C.F. (2001) 'Role functions, mechanisms, and hierarchy', in *Philosophy of Science*: 53–74.
- [9] Davidson, D. (2001) 'Mental events', in *Essays on actions and events: Philosophical essays Vol. 1*. Oxford: Oxford University Press.
- [10] Dennett, D. (1991) *Consciousness Explained*. London: Little, Brown.
- [11] Fiala B.J. (2012) 'Toward a Debunking Account of the 'Explanatory Gap''. Available at http://www.u.arizona.edu/~fiala/files/Fiala_WritingSample.pdf.
- [12] Fodor, J. (1974) 'Special sciences and the disunity of science as a working hypothesis', in *Synthese* 28: 97–115.
- [13] Hardin, C.L. (1987) 'Qualia and materialism: Closing the explanatory gap', in *Philosophy and Phenomenological Research* 48/2: 281–298.
- [14] Harman, G. (2007) 'Explaining an Explanatory Gap', in *APA Newsletter on Philosophy and Computers* 6/2.
- [15] Jackson, F. (1982) 'Epiphenomenal Qualia', in *The Philosophical Quarterly* 32: 127–136.
- [16] — (1986) 'What Mary didn't know' in *The Journal of Philosophy* 83/5: 291–295.
- [17] Kostić, D. (2011) 'The explanatory gap account and intelligibility of explanation', in *Theoria*, Beograd 54/3: 27–42.

- [18] Kripke, S. A. (1980) *Naming and Necessity*. Harvard: Harvard University Press.
- [19] Levine, J. (1983) 'Materialism and qualia: The explanatory gap', in *Pacific Philosophical Quarterly*, 64/4: 354–361.
- [20] — (1993) 'On leaving out what it's like' in M. Davies & G. Humphreys (eds.) *Consciousness: Psychological and Philosophical Essays*. Oxford: Blackwell Publishers.
- [21] — (2001) *Purple Haze: The Puzzle of Conscious Experience*. Cambridge, MA: MIT Press.
- [22] Nagel, T. (1974) 'What is it like to be a bat?', in *The Philosophical Review* 83/4: 435–450.
- [23] McGinn, C. (1989) 'Can We Solve the Mind–Body Problem?', in *Mind* 97/391: 349–366.
- [24] Taylor, E. (2015) 'Explanation and the Explanatory Gap', in *Acta Analytica* 31/1: 77–88.
- [25] Tye, M. (1999) 'Phenomenal consciousness: The explanatory gap as a cognitive illusion', in *Mind* 108/432: 705–725.

Conceiving ‘Ideally’ and its implications for the Modal Argument*

Ahmad Jabbar

University of Manchester

Introduction

In this paper, I argue that Kripke’s modal argument does not suffice to refute identity theory. In the modal argument, Kripke relies majorly on the premise of conceivability’s entailing possibility. My attack is specifically geared towards this premise. I consider a famous example in which conceivability does not entail possibility. I then consider Chalmers’ distinction between *prima facie* and ideal conceivability. Under Chalmers’ distinction, conceiving of minds as existing independently of bodies is distinct from the conception in the example presented, in that the conceivability in the former is ideal and in the latter is *prima facie*. However, I present a major objection to Chalmers’ distinction via a thought experiment. I show that Chalmers’ version of ideal conceivability is not intuitive. I then present my own distinction between *prima facie* and ideal conceivability. If Kripke claims the conception of mind existing independently from bodies to be ideal under my distinction, then Kripke’s argument is highly trivial. If he does not, then his conception is of the same type as the one in the example discussed.

In the first section of the paper, I will briefly explain identity theory. Furthermore, I will explain Kripke’s modal argument and show how it poses a threat to identity theory. In the second section, I will highlight how Kripke relies substantially on the premise involving the principle of conceivability’s entailing possibility. My discussion in this paper will concern itself only with the aforementioned premise. I will answer a possible objection that can be extracted from Chalmers’ discussion of the distinction between *prima facie* and ideal conceivability. In the last section, I will present my own distinction between the two forms of conceivability and show how Kripke’s conceivability is not ideal. Any attempt at making it ideal would make the conclusion of Kripke’s argument highly trivial.

*Delivered at the BJUPS Winter Conference, 14–15 January 2017 at UCL.

Identity Theory

Identity theory claims that our mental processes are type identical to the psycho-neural goings-on in our bodies. It relies on the contemporary scientific findings which relate each mental event with a certain neural state. To use Kim's example, as electric discharge from the clouds and lightning are numerically identical phenomena, pain and C-fibre firings are identical to each other.

The Modal Argument

Kripke, however, denies that our mental events are type identical to certain physical goings-on in our bodies. Before presenting the argument, it would be better to shed some light on the concept of rigid designation. There are certain referring expressions like 'the philosopher who wrote *Word and Object*'. This expression does not necessarily refer to Quine. This is so because, in a possible world, someone else has written the book. Therefore, the expression 'the philosopher who wrote *Word and Object*' does not refer to the same object in all possible worlds. On the contrary, an expression like 'Quine' is a proper name. For it to refer to another object in a possible world would be impossible. Therefore, it refers to the same object in all possible worlds and is a rigid designator.

Similarly, a mental event like pain would also be a rigid designator. Kim offers us a way to look at the rigid designation of 'pain'. "A world in which nothing ever hurts is a world without pain."¹ The physical process with which identity theorists identify pain is C-fibre firing. It is true of C-fibre firings as well that we cannot identify them with an event which is not the excitation of our C-fibres. It is the excitation of C-fibre firings that defines for us an event like C-fibre firings.

Hitherto, we have shown how natural kind terms like 'pain' and 'C-fibre firings' are rigid designators. Kripke² uses the concept of rigid designation of mental and physical events to refute identity theory. If being a rigid designator means to refer to the same object in each and every possible world, and pain is identical to C-fibre firings, then 'pain' and 'C-fibres firings' refer to the same object in every possible world. In other words, the identity between the

¹Kim [3] p. 120.

²Kripke [5] pp. 106–155.

referents of the two rigid designators holds in every possible world. Kripke thinks that the ‘illusion of contingency’ does not work in the case of ‘pain’. The illusion of contingency can be explained by an example. It might be true that heat sensations are not identical to molecular motion whilst heat is. What this means is that it is true that different organisms might react differently to the molecular motion but what they would be responding to would be the same thing, i.e. heat. However, in the case of pain, there are no transcendental pain sensations that are identified with C-fibre firings, as pain sensations are nothing but pain itself. At first, we might be tricked into thinking that, just as different organisms have different responses to molecular motion, they would have different responses to pain as well. However, to have a response to or to sense pain is what pain itself is. Therefore, the identity is not contingent and has to be necessary.

Another way to elucidate the premises above is as follows:

1. Let us assume that pain is identical to C-fibre firings.
2. Leibniz’s Law Application: If pain and C-fibre firings are identical to each other, then whatever is true of pain is true of C-fibre firings.

(P1) It is necessarily true of pain that it is identical to pain.

(P2) If it is necessarily true of pain that it is identical to pain, and Leibniz’s Law holds, then it is necessarily true of C-fibre firing that it is identical to pain.

The third premise that Kripke introduces is that it is conceivable that his mind exists without his body. He assumes the principle of conceivability’s entailing possibility, and so purportedly shows that it is true in at least one possible world that his mind exists without his body. However, if that is true then it is not necessarily true that our minds are identical with our bodies. This shows how the property of being identical to pain that is true of pain is not true of C-fibre firings. Therefore, pain and C-fibre firings are not identical in any possible world. In other words, if an identity between the referents of two rigid designators does not hold in one possible world, it does not hold in any world.

1 ‘Ideal’ Conceivability

It might at first sight seem as if the argument presented by Kripke is impenetrable. However, it should be noted that to prove the negation of the identity in a possible world, Kripke is relying on *a priori* reasoning. It is not the case that he has access to all worlds. Consequently, he is forced to use the principle

of conceivability's entailing possibility. It is important here to clarify that my discussion does not rely on the 'illusion of contingency' strategy since this is not needed. If I only prove how conceivability does not entail possibility, then Kripke's argument stands on a mere assertion.

There is not much need to explain the principle as it is self-evident. If I can conceive of a proposition X to be true, then X is true in a possible world. I can conceive of pegasus. Therefore, in a possible world, there are pegasuses.

Kneale presents an example of the Goldbach's conjecture to deny conceivability's entailing possibility.³ Goldbach's conjecture is an unproven mathematical conjecture. In virtue of being a mathematical proposition, if it gets proven or falsified, it would be true or false respectively in all possible worlds. However, I can conceive of its being proven and I can also conceive of its negation. Considering that it is either necessarily false or necessarily true, I would be conceiving of something which is impossible. Conceivability in this scenario would not entail possibility. Therefore, conceivability is not the right guide for inferring what is possible.

Chalmers, however, argues that the way we conceive of Goldbach's conjecture's being either true or false is not how 'ideally' things should be conceived for it to entail possibility.⁴ He expounds upon various forms of conceivability and says that our conception regarding the truth or falsity of a yet-to-be-proven conjecture is *prima facie*. Let us first understand what *prima facie* means. If a situation is *prima facie* conceivable for a subject, then the subject has conceived of that situation on little reflection. An example of such conceivability would be for me to conceive of a right-angled triangle as having squares of two sides (opposite and adjacent) not equal to the square of its hypotenuse. In this conception, I have not considered the nature of a right-angled triangle. The nature of right-angled triangles is such that it precludes the possibility of this conception. My conception was made capriciously or on little reflection.

Some familiar purported counter-examples to the claim that conceivability entails possibility are really counter-examples to the claim that *prima facie* conceivability entails possibility.⁵

Chalmers thinks that the conception of neither Goldbach's conjecture nor its negation involves ideal conceivability. Therefore, this example does not

³Kneale [4] pp. 79–80.

⁴Chalmers [2] p. 31.

⁵Ibid. p. 316.

counter ideal conceivability as a guide to possibility. Kripke's argument remains intact so long as his conceivability is not *prima facie* but ideal under Chalmers' distinction of *prima facie* and ideal. In order to refute the modal argument, I need to show either of the two following things. I need to either show how my conception of Goldbach's conjecture is ideal or how Chalmers' distinction is flawed.

I concede the fact that the Goldbach's conjecture example does not involve ideal conceivability. However, Chalmers has not substantially explained the distinction between *prima facie* and ideal conceivability. He characterizes ideal conceivability as conceivability involving ideal rational reflection. I will argue that Chalmers has equivocated the notion of ideal conceivability. Once I expound upon the equivocation, it would be easier to see how Kripke's conception of his mind's being independent of his body is not ideal. Therefore, it cannot guide us to a possible scenario.

So, what is it that Chalmers says to distinguish ideal conceivability from *prima facie* conceivability?

S will be ideally conceivable when ideal rational reflection detects no contradiction in the hypothesis expressed by S.⁶

My objection to this view is that it is quite possible that a scenario is conceivable without any apparent contradictions; however, it still is not 'ideal' in the intuitive sense. Chalmers' use of the word 'detects' alludes to the epistemological background of the conceiver. If our epistemological background regarding a particular conceived scenario is imperfect, we cannot be sure whether there is some 'not-yet-found contradiction'. In other words, we cannot 'detect' the contradiction. A scenario would still be possible according to ideal conceivability according to Chalmers if we are not aware of a latent *a posteriori* contradiction.

This can be shown via the following hypothetical scenario:

1. X has empirical knowledge regarding only two entities: *p* and *q*.
2. X believes *a priori* that *p* and *q* can occur together without any contradiction.
3. X does not know that *r* exists, and whenever *p* occurs, *r* accompanies it.
4. X does not know that *r* and *q* cannot occur simultaneously.

⁶Chalmers [1] p. 146.

5. X can conceive that p occurs \wedge q occurs without ‘detecting any contradiction in her hypothesis’.
6. However, we know that p occurs \wedge q occurs entails a contradiction i.e. $(p$ occurs \wedge r occurs) \wedge q occurs.

It is apparent from the scenario above that Chalmers’ condition of not detected contradiction for his version of ideal conceivability is met. However, X’s conception of the situation p occurs \wedge q occurs is not intuitively ideal as it involves an undetected contradiction. X can conceive of the situation without detecting any contradiction. However, the scenario does involve a contradiction and is consequently impossible. Chalmers’ version of ideal conceivability here, is entailing a scenario that is impossible. Therefore, Chalmers’ use of the word ‘detects’ makes his distinction vacuous.

2 J-Ideal Conceivability

It is quite evident now that neither *prima facie* nor Chalmers’ version of ideal conceivability can be considered a guide to possible scenarios. Thereby, in this section, I will present my own distinction between *prima facie* and ideal conceivability. Henceforth, my use of the word ‘ideal’ would not be in Chalmers’ sense. Moreover, it will be shown how Kripke’s use of conceivability in his modal argument is not ideal. If I show his use not to be ideal, it cannot guide us to a possible scenario. Then his argument would have no force left.

Let us assume that there lives an isolated family in Antarctica consisting of a couple and a kid named X. A new baby is born, and the couple tells their first child X that her sister descended from the skies. For X, it would be conceivable for her sister to counterfactually be born to some other couple. If we were to apply the principle of conceivability’s entailing possibility here, then it is possible that X’s sister was born to another couple (with the same genetic profile she actually has).

However, we know that such a thing is impossible. This is so because we are acquainted with biological sciences, and we know that a child inherits its parents’ genes. Therefore, it is neither conceivable nor possible that a child is born to different parents. A child born to different parents would not be identical with X’s sister.

This shows that what we conceive of is dependent on our empirical knowledge regarding the scenario at the time of conception. Therefore, conceiving of a scenario regarding which we have incomplete knowledge would not be ideal conceivability. Ideal conceivability, under my version, is distinct from

prima facie conceivability in that it requires complete knowledge regarding the entities featured in our scenario. I will call this sort of ideal conceivability, J-ideal conceivability.

So long as in a conceived scenario where there is a possibility of it containing a detectable contradiction about which we would get to learn in the future, the conceived scenario cannot be a guide to a possible scenario. Let us rephrase our distinction between *prima facie* and J-ideal conceivability now. J-ideal conceivability differs from *prima facie* conceivability in that it requires the preclusion of any potentially detectable contradiction in the conceived scenario.

Let us now consider Kripke's conception of his mind existing independently of his body. We do not know what pain is. We do not know whether it is identical to C-fibre firings or not. If in the future, natural sciences find pain and C-fibre firings to be identical with each other, to conceive them occurring independently would entail a contradiction. That would not be 'ideal' under Chalmers' version as well at some time in the future when we have empirical knowledge regarding the identity. It is fully possible that they might not be identical. In that case, the conception would be possible. However, we have incomplete knowledge. To conceive of pain and C-fibre firings to exist independently at this moment is to exclude the possibility that natural sciences might find them to be identical in the future. Therefore, we can conclude that Kripke's conceivability is not J-ideal and consequently does not entail the possibility of his conceived scenario.

Kripke's motive for his modal argument is to elucidate the nature of pain. For this purpose, he employs the 'conceivability entailing possibility' principle. He can conceive of pain's existing without C-fibre firings. From the discussion above, it would be clear that his use of conceivability is not ideal in the intuitive sense. If Kripke claims it to be J-ideal, then there is a bigger problem for him. J-ideal conceivability would involve exclusion of the fact that pain can be identical to C-fibre firings in the future. By excluding this possibility, Kripke would be making a claim regarding the nature of pain which is something he is trying to elucidate through his argument.

For Kripke's conceivability argument to work, pain must not be identical to C-fibre firings. Only then can Kripke's conceivability be J-ideal, as he would be ruling out a possible contradiction. By virtue of conceiving pain independently of C-fibre firings, Kripke is excluding the possibility of pain being identical to C-fibre firings in all possible worlds.

There is a possibility of pain being identical to C-fibre firings in all possible worlds. If that is the case, then Kripke's conception would be entailing a contradiction. It would be analogous to the conception of molecular mo-

tion's being non-identical to heat. Conceiving of molecular motion's being non-identical to heat does not entail its possibility, as molecular motion is necessarily identical to heat. Under J-ideal conceivability, to conceive that pain can exist independently of C-fibre firings is to assume that one has knowledge that pain cannot be identical to C-fibre firings. This makes Kripke's modal argument highly trivial. The conclusion of the modal argument is that pain is not identical to C-fibre firings. Only once that conclusion is known can one conceive of pain as existing independently of C-fibre firings. To prove that pain is not identical to C-fibre firings in the conclusion, Kripke would be implicitly assuming that pain cannot be identical to C-fibre firings. Therefore, Kripke's argument becomes circular; at best an argument with a conclusion that is entailed trivially by its premises.

In this paper, I have shown how Chalmers' version of ideal conceivability is too weak. This was shown by a hypothetical scenario. I presented my own distinction between *prima facie* and ideal conceivability. Consequently, I showed that Kripke's conceivability can only run as a very weak version of ideal conceivability which does not entail possibility. Lastly, I showed that if Kripke's conceivability subscribes to a more intuitive version of ideal conceivability, then the conclusion of his argument is only entailed trivially by the premises.

References

- [1] Chalmers, D.J. (2002) 'Does Conceivability Entail Possibility?', in T. Gendler & J. Hawthorne (eds), *Conceivability and Possibility*. Oxford: Oxford University Press.
- [2] — (2009) 'The Two-Dimensional Argument Against Materialism', in A. Beckermann, B.P. McLaughlin & S. Walter (eds), *The Oxford Handbook of Philosophy of Mind*. New York: Oxford University Press.
- [3] Kim, J. (2011) *Philosophy of Mind*. Boulder, CO: Westview Press.
- [4] Kneale, W.C. (1949) *Probability and Induction*. Oxford: Clarendon Press.
- [5] Kripke, S. (1980) *Naming and Necessity*. Cambridge, Mass: Harvard University Press.

Against Atomism*

Gareth Pearce

University of York

I will first explain what I mean by atomism, then show how it fails to explain two types of observable phenomena: repulsive forces (e.g. same-charge magnetism) and attractive forces (e.g. gravity and opposite-charge magnetism). There is a common third objection to atomism, that it doesn't successfully account for phenomenal experience: this is, however, well-trodden philosophical ground. As such, I would only succeed in reproducing the arguments of others, rather than producing anything new myself. I will also assume laws of physics that are more or less Newtonian in nature.

1 What is Atomism?

I will take atomism to mean the following.

- (1) The only existing thing is matter; between matter is void.¹
- (2) All matter is made up of irreducible simple parts called 'atoms'. Complex structure comes from the arrangement of atoms or their movement in a system.²
- (3) Atoms have the properties of shape or size, mass, motion and location (spatial and temporal). They have no other properties.³
- (4) Matter can neither be created nor destroyed.⁴

Whilst I have drawn this account of atomism from Epicurus and Lucretius, I do not wish to limit atomism to precisely what was said by him, or any other atomist. I am more interested in the sorts of moves that an atomist could make, rather than any particular moves Epicurus actually did make.

*Delivered at the BJUPS Winter Conference, 14–15 January 2017 at UCL.

¹Epicurus [1] p. 39.

²Ibid. pp. 40–41.

³Ibid. pp. 43, 54.

⁴Ibid. p. 38.

2 Our *prima facie* concern with atomism

If atomism is true then we would expect the complex structures formed by the combination of atoms to also only exhibit the properties discussed in (3). No additional primary property can arise from the arrangement of atoms in any interesting way. For instance, putting the shape of two atoms together just makes a bigger shape; resolving the motion of two atoms in a collision will only lead to more motion. Equally, the properties could not suddenly emerge as a feature of higher-order structure as there would need to be some object to cause this property to come into being: that object could only comprise atoms themselves as atoms are the only things that exist but atoms lack any property of the capacity to do this. So how then do we explain the plethora of other properties that the objects we encounter every day seem to have?

The reply is an obvious one: all other properties we see (for instance the sky's being blue) are formed by complex arrangements of atoms. For instance, to say "the sky is blue" means "the sky is omitting light particles at a frequency of 650 THz".

This has the possibility to turn into a real strength of the atomist theory. If atomism can explain every macroscopic property we encounter in terms of the movement and arrangement of atoms, then we would have a theory that posits a very small ontology with no metaphysically odd objects and total explanatory power. For this to be the case, however, atomism must explain all properties we encounter. I will show that atomism fails to explain repulsive forces at a distance (RF) and attractive forces at a distance (AF) and thus does not have total explanatory power.

3 Repulsive forces

Repulsive force (RF) is where we see two particles (A and B) at a distance appear to 'push away' from each other. Ordinarily, this would be explained through a non-corporeal force between the particles. However, this solution is not accessible to atomists as force is not a property of atoms mentioned in (3). So atomism must explain RF through the movement of particles. As A and B aren't touching, there must be a series of 'carrier' particles (P) that obey atomism and act on A and B to create the impression of RF (in much the same way light particles create the impression of blue).

Simply, P could leave A in the direction of B, causing A to move away from a stationary B as P 'pushes off'. P would then hit B and cause B to move away from A. Three immediate issues arise in relation to repellent forces.

First, there could be no neutral particles (e.g. a neutron if RF was magnetism) as any object hit by P would gain momentum. This is easily solved, neutral particles could be mesh-like in shape and allow the carrier-particle to pass through unrestricted. Second, given that A and B are type-identical particles why does particle P leave A but a similar particle P^* not leave B as well?

Third, P leaves A before it hits B so A would move before B. Most instances of RF we observe are instantaneous - even across great distances (so P can't just be really fast and give the impression of immediacy when, in fact, A moves an instant before B).

The latter problems can be solved together by adjusting our mechanics of RF. A particle P_1 can leave A and a particle P_2 can leave B (the two P particles 'pushing off' causing the motion of A and B). P_1 and P_2 then collide and coalesce with their resultant momentum being 0, while the combined P_{1-2} remains stationary.

But what would cause A and B to release P at the same time and at the right time? If everything can be explained through size, motion and position there cannot be any mystical 'communication' between the particles. Atomism couldn't postulate a 'communication particle' as we'd need an explanation of what causes the 'communication particle' to be released at the right time. The timely release of which, under atomism, would have to be explained by a 'communication communication particle' and so on *ad infinitum*. So if RF is explainable through atomism and there can be no 'communication' between A and B, then A and B must be emitting P particles all the time (or at least on regular intervals).

Our mechanic for RF would work as follows: A and B emit a flow of P from all angles (there would be no net motion at this point as the 'push off' from P particles on opposing sides of A/B would cancel each other out). P particles travelling from A towards B and vice versa would collide between A and B, rebound (we would have to change the 'shape' of our P particles so they rebound rather than coalesce) and return back to A and B, colliding with them and causing their motion.

But a constant flow of P particles would need an explanation of where the P particles comes from. If we accept (4) then they cannot be created ex nihilo. Therefore, particles like A and B would need a resupply method or run out of particles. We could amend atomism and remove (4) or we could find some way of explaining how A (or B) can 'collect' P particles without it affecting our description of RF.

The first route would require an explanation of how P particles just happen to

appear in such a way that they facilitate RF between A. This couldn't be because of any property of A that causes P particles to appear in the appropriate way inside them as this would contravene (3). Nor could they appear randomly with no momentum to be 'collected' by A. The P particles would lack the momentum needed to move out of A and create RF. Nor could they simply appear in a random position with momentum as they would then transfer momentum to A when being 'collected' and cause totally random movements in A that we do not observe; unless, that is, there were multiple P particles colliding with A and cancelling each other out.

For this to be possible, there would either need to be a substantial quantity of P particles being created to balance out the net force on A (with small random variations) or P particles being created in pairs with one on either side of A. If the former then the universe would soon become swamped with P particles that would interfere with other interactions between particles and, as the total number of P particles increases, the universe would become swamped and would lose form as the tide of P particles pushes everything apart. If the latter, then one would need to postulate a property of A that causes P particles to appear on either side of it, moving directly towards its centre (were they to appear off centre A would gain momentum from the 'pinching' effect of two P particles) and this would contravene (3).

Our other option for the 'resupply problem' was to have A collecting P particles in some way. Were lots of P particles flying about and were they to rebound and transfer momentum when colliding with A (the movement of A wouldn't really be affected by this on its own as collisions on opposing sides would cancel each other out), then when particles like A and B come close to each other the P particles rebounding from A towards B and vis versa would rebound off each other in the middle and oscillate, colliding with A/B and the opposing P particles in turn. This would cause a net force on A and B away from the oscillation explaining RF.

However a single rogue P particle, that just happened to move between A and B, would disrupt the oscillation by knocking the P particles between A and B off course. Given that we've postulated a large number of P particles such interference would be almost inevitable and nearly constant allowing for no constant oscillation of P particles, no transfer of momentum and no RF. Moreover were A to have any velocity itself, then it would experience more P particle collisions on its forward moving face than on its back face and hence would have a 'drag' effect. So any particle interacting with P, like A or B, would always decelerate to a stop, and this is not observed.

Therefore there appears to be no obvious way that atomism can explain RF.

4 Attractive Forces

Attractive force (AF) is when we see two particles A and B appear to ‘pull’ towards each other at a distance (e.g. gravity). In the same way that atomism requires force carriers of RF, so too does it require force carriers of AF (call these ‘Q particles’ to avoid confusion). An atomist account of AF will face many of the same challenges as RF. For instance, an inability of A and B to ‘communicate’ requires a constant flow of Q particles from A and B which in turn leads to the same ‘resupply problem’ for A and B. I will not repeat this ground as the argument is identical.

AF faces an additional problem that RF does not. Whilst providing an exact mechanic of RF is proving difficult, the notion of how a carrier particle P’s ‘pushing off’ from A and then colliding with B would lead to a net repulsion between A and B is a solid one. The same is not true for AF as it has the added difficulty of explaining how a Q particle can ‘push off’ from A and collide with B and yet cause them to move closer together. I will explore two possible solutions to this problem: negative particles and boomerang particles. I’ll show how both of these fail.

First of all, negative particles. An obvious solution to the problem would be to let Q particles be ‘negative’ in mass and hence a collision with a Q particle would cause an object to move towards the collision rather than away from it. *Prima facie* this would contravene (3) and so would constitute a rejection of atomism. However, I don’t think this is strictly the case: (3) postulates that the only properties of an atom are size, mass, position and speed but does not say that mass must only be positive. However, I think to accept the notion of negative mass would require a very strange understanding of what we mean by “mass”. Mass is usually understood to be the amount of matter in an object; whilst it’s possible to conceive of different densities of matter in atoms, it is hard to see how that could ever be negative as this would require, not just an absence of matter but a negative amount of it. That said, it’s not an obviously contradictory idea and it could certainly be mathematically modelled so I think it is a move that’s consistent with atomism. Two issues, however: atomism would lose some of the metaphysical simplicity that makes it otherwise so appealing. Secondly, the exchange of negative particles is still going to encounter the sorts of problems of communication and supply that RF does.

Our second solution was ‘boomerang particles’. The idea of a boomerang particle is that a Q particle could ‘push off’ from the side of A facing away from B, loop behind B and collide into B from the other side. The net force on A would be towards B and on B towards A so the two particles would move to-

gether. B, being the same type of particle as A, would have to behave in the same way, so a Q particle would also leave B on the side away from A, loop round, and hit A on the far side.

There are, however two, problems. first, the motion of Q particles would need to be curved in order to loop round from one particle to the other. To explain this, either Q particles must have different laws of motion applying to them than the other particles in the universe or Q particles must be giving off smaller particles in order to readjust their motion. If the former then proponents of atomism would need to explain why the laws of motion arbitrarily affect one particle in a different way to all the others. If the latter then Q particles are going to suffer from ‘resupply’ problems.

The second problem is that for AF to take place, the Q particles would need to curve in exactly the right way between A and B, and this is dependent on the distance between A and B, which could vary. So atomism would need to explain why the Q particles would happen to curve in sooner when A and B are closer and why they would curve in later when A and B are further apart. Here there could be some kind of communication particle (C particles) that could control this. If A and B were emitting C particles (that have negligible weight so don’t affect the motion of A or B) that were then colliding with the opposite particle (e.g. C particles from A colliding with B) and bouncing back to the particle they came from, then the time period of the oscillation would be a function of the distance between A and B. If the angle that Q particles were released from was a function of the frequency of impacts of C particles from the opposite side of A/B (e.g. A/B could have ‘fins’ that control the direction of Q particles and change angle as a result of C particle collisions) then this could explain how the Q particles happen to curve in exactly the right way.

However it would be an astounding coincidence in the fundamental nature of reality that the rate of C particle collisions caused by an oscillation of C particles of a certain distance would cause the Q particles released to curve across that same particular distance. Moreover the exchange of C particles would be difficult to explain if A and B have any relative motion (other than that caused by AF between them). If, for instance, A and B were moving towards each other, then the rate of C particle collisions would be greater than would otherwise be and the Q particles should fall short of the target and AF should not take place. Or if A and B were moving towards one another but set to cross without touching then the Q particles would need to bend differently in order to hit A and B in the right way. So C particle exchange does not seem to be a sufficient explanation for the correct amount of curvature in Q particles.

The ‘correct curvature problem’ could be solved by having Q particles com-

ing from A and B all at different angles such that, no matter what distance is between A and B, at least one Q particle will find the right spot. First of all this would massively worsen the re-supply problem for A and B. Secondly, the Q particles coming out of B that are taking a shorter, more angular trajectory would collide with the Q particles coming in from A on a longer trajectory, preventing the collision that causes AF. Thirdly, even if the Q particles could reach B from A and vice versa, It would be hard to describe a picture that explained AF as we observe it. We would see a number of Q particles hitting B/A on one side and not on the other. For instance If A is to the left of B and Q particles curve in a clockwise arc, then there would be more collisions in the top right side of B than anywhere else on B and more on the bottom left side of A than anywhere else on A. This would cause A and B to spiral into each other rather than simply pull towards each other as we see in AF.

So overall boomerang Q particles do not seem to effectively explain AF. Therefore there appears to be no obvious way that atomism can explain AF.

5 Conclusion

I have shown how atomism comes into considerable trouble explaining attractive and repulsive forces at a distance and, barring a novel solution from atomists that I have not considered, lacks complete explanatory power. No novel solution is forthcoming.

References

- [1] Epicurus (1994) *Epicurus: The Epicurus Reader*, B. Inwood & L.P. Gerson (trans.) Indianapolis, IN: Hackett Publishing Company.

Post-hoc* Memory Erasure as a Response to the Kane–Widerker Objection

Harry Sherwood

New College of the Humanities

The link between freedom and alternate possibilities seems a strong one. Suppose that an agent must decide between the mutually exclusive options of (A) crossing the road or (B) staying stationary at a time t . It seems reasonable that the agent's choice is a free one only if both options were possible at time t . In other words, even though the agent in fact chose to cross the road, they *could have* remained stationary. One could argue that these 'possible alternatives' AP are *necessary* for the choice to be free, but some such as Frankfurt have challenged this claim, citing examples where an agent is unable to choose otherwise than he did and yet appears to have chosen freely. In these counterexamples, freedom is present without AP and so the necessary link is severed. My aim in this essay is to show that the Frankfortian counterexample is faulty and that, consequently, the principle of alternate possibilities (PAP) is still necessary condition for freedom of the will. In the first part, I introduce the characterisation of alternate possibilities as requisite for free will before summarising Frankfurt's objections to it. Next, I discuss the epistemic problem raised by Kane and Widerker: the actions of a 'counterfactual intervener' in (PAP) counterexamples are only convincing in a deterministic universe where prior prediction is possible.¹ If it isn't possible to predict an event before it occurs (for example in an indeterministic universe) then Frankfurt's counterexamples lose their traction. I try to rescue Frankfurt by modifying his objection so that it doesn't require any prediction, but can this survive the criticisms which I anticipate of it? If Frankfurt's objection cannot be rescued, (PAP) survives. As a result, the possibility of choosing (B) is not necessary for one to have freely chosen (A).

Although Frankfurt considers the principle of alternate possibilities (PAP) mainly in terms of moral responsibility, it is possible to disentangle the moral component from the metaphysical one and talk about freedom as 'the ability to do otherwise'. The following is often said.

*Delivered at the BJUPS Winter Conference, 14–15 January 2017 at UCL.

¹Kane [3], Widerker [5].

(FRAP) An agent chooses freely only if said agent could have chosen otherwise.^{2 3}

Intuitions *supporting* FRAP can be clarified with the use of examples. For instance:

(A) **Context** Suppose that Billy is walking along the street when he is kidnapped, drugged, and taken to a warehouse. Bertrand, a malicious neuroscientist-turned-killer who wants another hostage dead (but is excellent at self-exculpation) has devised a small implant which he places into Billy's brain while he is unconscious. This allows Bertrand to implant compulsions into Billy, which Billy acts upon.

Action Billy awakens with the unquenchable urge to pick up a gun and shoot the hostage. It is as if his actions have a sort of gravity to them which he cannot resist. He picks up the gun and shoots the hostage.

What seems clear is that Billy was severely coerced into shooting the hostage. No amount of effort or willpower could have prevented his action, and so it is safe to conclude that Billy could not have done otherwise. Frankfurt provides similar examples of coercion in *Alternate possibilities and moral responsibility*, before providing a counterexample, which I paraphrase below.

(B) **Context** Billy is actually a murderer at heart, although he is integrated into society. Bertrand, who wants the hostage dead but would rather not show his hand, has set up the implant as in the previous scenario. Billy awakens, and, being the murderer that he is, looks upon the pleading hostage and begins to decide whether to kill her. If Billy does decide to shoot the hostage, the implant simply monitors the corresponding indicative neuronal pattern without intervening in his brain processes; however, if Bertrand's creation detects that Billy is about to choose *not* to shoot the hostage, it registers some corresponding neuronal pattern in his brain. This pattern triggers the device to intervene causally in Billy's brain processes, planting in him the same compulsion to shoot the hostage as in the case of severe coercion.

²"A person has free will if he is often in the position of having to decide which of two or more mutually exclusive courses of action he will follow, where each is such that he can follow it, or has the ability to follow it, or has the capacity or power to follow it should he so intend." (Van Inwagen [4])

³At this point in the essay, I will be focusing on the decision as opposed to the action, although perhaps a decision is merely a basic type of action in a Davidsonian sense.

Action To Bertrand’s disappointment (he had spent a whole day creating the implant), Billy wakes up and forms the completely self-authored decision to shoot the hostage. The implant merely monitors.

Let us consider the implications of Billy’s action(s) for FRAP. Firstly (I), it is clear that he could not have acted or decided otherwise than he in fact did: he literally had no choice in the matter; for as soon as the implant detected the telltale signs of a decision not to kill the hostage, it intervened. He could not have even decided otherwise, let alone acted on that decision. Secondly (II), since he did in fact kill the hostage of his own volition, he seems to have decided freely: *his* thoughts were the only ones present in the causal chain which blossomed into a decision. Yet if (I) and (II) are the case, then it is not true that FRAP, or that “an agent chooses freely only if said agent could have chosen otherwise”. By way of analogy, just as the existence of an oxygen atom bonded to two hydrogen atoms is sufficient for water to exist but does not cause or explain ‘water’, there is a sense in which having no alternative to a given decision is a *sufficient condition* for an agent to make that decision, even though the lack of an alternative may have played no causal or explanatory role in why they decided as they did.

Frankfurt’s objection to FRAP is esteemed by many such as Dennett as one of the great breakthroughs of compatibilism (the position where free will is said to be compatible with the truth of determinism). It would therefore be ironic to discover that they *require* the truth of determinism to operate. If this was the case, the objection would simply be begging the question with respect to compatibilism. This is the epistemic problem raised by Kane and later Widerker: in FRAP counterexamples, the counterfactual intervener (i.e Bertrand’s implant) can only be ‘counterfactual’ in a deterministic, predictable universe. Because the indeterminist libertarian does not accept such ‘foreknowledge’, the objection becomes implausible. This has the impression of being a potent criticism of Frankfurt’s objection, but I would like to express it in a way that makes it more clear.

What is the mechanism by which the device inside Billy’s head operates? As discussed earlier, it monitors neuronal activity and so can ascertain with a high degree of probability what Billy will decide, before he has even finished deciding. In this sense, it is a decision prediction machine, capable of combining information about the neuronal pattern at a point in time with information about the relationship between certain patterns of activity and decisions. But what is this ‘relationship’? Perhaps we can characterise the relationship as the arrow in (R) below:

(R) Initial Neuronal Pattern (INP) → Billy’s decision (D)

Such that Billy's decision (D) will be:

(D) INP + Relationship between past INP and (D)

Yet this bears striking resemblance to a simple form of nomological deduction formulated by Hempel, where an *explanandum* event (E) is explained in terms of a set of initial facts and covering or 'transportative' laws which bring them to match the *explanandum* event.⁴ For example:

(E) Initial facts + covering laws

If Bertrand's device 'reasons' or operates in a similar way to (D), then it is actually enthymematic in that it has suppressed a premise regarding the uniformity of nature. Without wishing to delve too deeply into induction of the Humean sort, it is of course true that to use a past rule in the present situation is to assert that the rule still applies – that the same set of causes necessarily produce the same set of effects. So perhaps (D) should be written more explicitly.

(D') INP + (R) + 'uniformity of nature'

But if Billy's neuronal implant operates in this manner, is it not guilty of determinism? After all, what is determinism if not the idea that, approximately, each effect follows as a necessary consequence from a set of causes? As John Widerker writes, if the process of (D) is "grounded in some fact that is causally sufficient (in the circumstances) for [Billy's] decision"⁵ then it is impossible for Billy *not* to have decided otherwise than he did, because his Initial Neuronal Pattern *caused* him to decide (D). Yet if Billy *had* to decide the way that he did given his INP (i.e. he was not able to even *start* deciding not to shoot the hostage) then our intuitions that he freely chose to shoot her start to waver.⁶

The point that allows Frankfurt's objection is that (1) Billy chooses freely and that (2) he nevertheless was unable to decide otherwise. In spite of this, the very counterfactual mechanism that ensures (2) implicitly rules out (1). Thus the conjunction of (1) and (2) is not possible – at least in the way that Frankfurt characterises it. In this way, FRAP goes unchallenged. I'm now going to

⁴See Hempel [1] and Hempel & Oppenheim [2].

⁵ Widerker [5].

⁶For the determinist, of course, Frankfurt's objection is metaphysically possible. But for the libertarian who is an indeterminist, the objection isn't possible, and so poses no problem. Frankfurt would need to explain why he thinks that determinism is true, but then he would be left begging the question. He wants to show that there is no conflict between an action being both free and also restricted, but in assuming determinism in his example he has not shown this but merely asserted it.

try to rescue Frankfurt by modifying his objection so that it does not require predictions founded upon determinism. Imagine a similar situation to the one with Billy and the hostage, but with the following differences:

Context Bertrand has realised that his predictions about what Billy will decide might be unreliable, so he designs a different device. The device monitors Billy’s neuronal patterns as before, but this time does not predict how he will decide. Instead, it simply waits until he *has* decided. The implant is configured such that, if Billy decides *not* to kill the hostage, his memory of having made the decision will be erased, and he will feel as if he is still deliberating.

Action Billy awakens as before, and sees the hostage. He wonders whether to kill her or not, and decides not to kill her. At this point, the device inside his brain whirrs into life, zapping specific areas of his hippocampus and prefrontal cortex and causing a sense of mild confusion as well as causing him to forget that he had made his mind up. Once again, he wonders whether to kill her, and, once again he decides against it. Yet Billy is, after all, a murderer – that’s why Bertrand kidnapped him in the first place: he theorized that, at least some of the time, Billy would freely decide to kill the hostage. After five iterations, Bertrand is about to give up hope; however, on this sixth one, Billy takes a little longer than before, and finally decides to pick up the gun and shoot the hostage.

Well – isn’t this interesting. It seems that the two conditions for Frankfurt’s objections are satisfied above: (1) Billy freely chose to kill the hostage and yet (2) he could not have decided otherwise. Neither is prediction needed or determinism assumed: the mechanism of the implant doesn’t involve any ‘foreknowledge’ or ‘prior sign’ of what Billy will decide and only interferes after the fact. Consequently, FRAP once again looks incorrect in the same way as Frankfurt originally intended because a case of free will without alternate possibilities is metaphysically possible.

I’d like to briefly anticipate and defend this revised Frankfurtian counterexample from criticism, starting with what I perceive to be the poorest. Firstly, one might ridicule the counterexample in terms of scientific plausibility, but there are three problems with this: (i) that such an implant and/or intervention is not particularly outlandish; (ii) that the technology required for the original Frankfurtian counterexample is actually *more* outlandish than the revised one; and finally that (iii) scientific plausibility is surely irrelevant in these thought experiments. I discuss (i)–(iii) below.

Regarding (i), there are two scientific or technological challenges for the sit-

uation in the counterexample to come to fruition. The first of these is the reading of what decision Billy has made. This reading could be made internally, through some complex analysis of brain function, or through more simple means. For example, the implant could be programmed to wait until Billy makes a clear physical manifestation of his decision, such as getting up to untie the hostage. The second (ii) technological hurdle concerns the implant's ability to induce specific decision-related amnesia. There is a good deal of evidence that specific neural pathways are responsible for the encoding of memories, and there already exist compounds that can ablate or destroy those pathways in mice. Supplanting a fully-fledged desire in Billy to kill the hostage seems like a much harder task, because desires are the sorts of things that are tightly connected to other beliefs and mental states which one would have to also alter. Frankfurt's original counterexample involves technology far more advanced than would be required in my counterexample. Finally, (iii) arguably in thought experiments the point is that the situations are *in principle* possible, or physically possible. There doesn't seem to be anything physically impossible about my counterexample.

Next, there is the criticism that this counterexample isn't really a fair treatment of the issues because it essentially winds back time repeatedly until poor Billy does as intended and so treats many decisions as if they are one decision, and many situations as if they are one. I don't agree: the counterexample does not involve the reversal of the flow of time or any such metaphysical oddity, but instead intervenes so as to affect the subjective experience of an agent. Neither does the counterexample erroneously conflate multiple decisions with one decision: the choices of Billy are still real and numerous, but Billy is made to re-decide multiple times.

Thirdly, one could assert that actually, Billy did decide otherwise, and so condition (2) of Frankfurt's objection is not satisfied. To say that he 'could not have decided otherwise' is simply false. I agree that this criticism is most pressing, and it should be developed. However, I would conclude thusly: there is a certain real sense in which Billy did not decide otherwise. All decisions to spare the hostage are erased, so is it fair to speak of them as existing?

References

- [1] Hempel, C. (1942) 'The Function of General Laws in History', in *Journal of Philosophy* 39: 35–48.

- [2] — & P. Oppenheim (1948) ‘Studies in the Logic of Explanation’, in *Philosophy of Science* 15: 135–175.
- [3] Kane, R. (2000) ‘The Dual Regress of Free Will and the Role of Alternative Possibilities’, in *Noûs* 34 /14: 57–79.
- [4] Van Inwagen, P. (1983) *An Essay on Free Will*. Oxford: Oxford University Press.
- [5] Widerker, D. (1995) ‘Libertarianism and Frankfurt’s attack on the principle-of-alternative-possibilities’, in *Philosophical Review* 104/2: 247–261.

Kant's Epistemology of Testimony*

Anni Raty

University of Cambridge

Suppose you tell me that the weather in London was dreadful last week. And suppose I come to believe that the weather in London must have been dreadful, since you say so. Is my believing that the weather was dreadful epistemically justified; is it a candidate for knowledge? It depends on the details. Most people would say: if I had good reason to think that you're not the kind of person to deceive others about these kinds of matters, and if I am confident that you are knowledgeable about the weather in London last week, then it's plausible that I can come to know that the weather was dreadful by taking your word for it. But if I have reason to think that you are a pathological liar, or that you suffer from persistent hallucinations of rain, then I shouldn't have believed what you said; I was being credulous and epistemically irresponsible in doing so.

The conditions under which a testimonially based belief is epistemically justified, or the conditions under which testimony can result in testimonial knowledge, make up the core subject matter of the epistemology of testimony. There are two eminent Enlightenment thinkers whose views on testimony have been carefully studied and whose ideas have greatly influenced contemporary authors on the topic: these are David Hume and Thomas Reid. On Hume's view, the acceptance of testimony must be backed up by inductive evidence for thinking that the speaker's report is knowledgeable and true; on Reid's view, no such evidence is necessary, but rather, we have an epistemic right to trust the testimony of others without having to acquire positive reasons to think that they are reliable. A third major Enlightenment thinker is sometimes said to have nothing much to say about testimony – this is Immanuel Kant. This paper argues that, to the contrary, Kant's writings contain a nuanced and attractive view of the epistemology of testimony on which we often are required to have positive reasons to think that our interlocutors are reliable, but on which we may *sometimes* be warranted in simply taking their word for what is said. It presents the first reading of Kant's views on testimony which does not represent his view as one on which we have an epistemic licence to trust people's say-so without positive evidence of their reliability.

The outline of the paper is as follows. Section 1 lays out Kant's general epis-

*Delivered at the BJUPS Winter Conference, 14–15 January 2017 at UCL.

temic theses about testimony, and considers what is meant by his claim that testimonial knowledge rests on the reliability of testimony. I will argue that the reliability of testimony as a source of knowledge depends on the uptake of testimony being combined with an assessment of the speakers' credibility. Section 2 examines Kant's account of what it takes for a testifier to be 'credible'; in sum, testifiers are credible when they know what they are talking about, and they are speaking truthfully. Section 3 begins an inquiry into the norms that must govern testimonial uptake, if testimony is to serve its function as a source of knowledge and justified belief. There are two norms that we can attribute to Kant on the basis of textual evidence: a moral principle that requires us to presume, until proven otherwise, that our speakers are speaking sincerely; and an epistemic principle that requires us to make a judgement about the epistemic credentials of our interlocutor before we believe things on their say-so. The final section argues that it is just these principles that are required for testimonial exchange to play a role that Kant intends it to play: for it to serve as a means to epistemic improvement, and as a corrective on our fallible understanding.

1 The reliability of testimony, the credibility of witnesses

The fact that Kant's epistemology of testimony has attracted little attention is not due to any lack of textual evidence for his views on the topic. Testimony is present through Kant's mature writings from the *Critique of Pure Reason* in 1781, to the *Logic* published in 1800. In addition to these, the 1786 essay *What Does It Mean to Orient Oneself in Thinking?* and the *Lectures on Logic*, collated from notes from lectures given between the 1770s and 1800, contain detailed examination of topics such as the credibility of witnesses, the nature and limits of testimonial knowledge, the role of other people's testimony as a means to our own epistemic improvement, and various moral aspects of testimony. Our inquiry here will begin from a concise but informative passage in the *Jäsche Logic*. The passage reads:

We can accept an empirical truth on the testimony of others with the same certainty as if we had attained it through *facta* of our own experience. In the former kind of empirical knowledge there is something deceptive, but also with the latter kind. Historical or mediate empirical knowledge rests on the reliability of testimony. The requirements of an irrefutable witness include *authen-*

ticity (competence) and *integrity*.¹

Here, Kant claims that a belief based on another person's testimony can amount to knowledge just as well as a belief based on first hand experience can, and that testimonial knowledge in some sense 'rests on' the reliability of testimony. If a testifier is to be believed, then she needs to have 'competence' and 'integrity'.²

The meaning of these last two requirements will be taken up in the next section. The first claim I want to argue for is that the reliability of testimony, which is the ability of testimony to serve as a source of epistemically justified belief for hearers, rests on the hearer's assessment of the credibility of witnesses. Consider the following remarks:

To attain proper cognition of the truth... it is required in any case that one also investigate whether someone speaks the truth or lies.⁴

Seeing belief is belief that is combined with examination of the witness and rests on this. Blind belief, however, is belief that accepts testimony without examination or investigation of its credibility. The former leads to truth, but the other is the path toward error, thus harmful, but the former is useful.⁵

¹Kant [6] JL 72n. See also VL pp. 896–898. All references are to the Cambridge editions of *Lectures on Logic* and *Critique of Pure Reason*. I abbreviate JL for *Jäsche Logic*, BL for *Blomberg Logic*, DWL for *Dohna-Wundlacken Logic*, and VL for *Vienna Logic*.

²In addition, Kant claims that only *empirical truths* can be known by testimony; what can be known *a priori* should never be accepted on other's say-so.³ I note this important aspect of Kant's account of testimony only in passing for the following reasons. First, there are no clearly expressed views in the literature as to why Kant restricts testimony to empirical propositions, and settling the issue from start is beyond the scope of this paper. Second, the implications of this restriction are striking: it seems to follow for example that mathematical and moral knowledge cannot be acquired by testimony. For these consequences alone, the issue deserves a lengthy discussion of its own. Third, the restriction invites questions about whether the degree of confidence warranted by testimonial evidence suffices, on Kant's view, for knowledge at all: some commentators have suggested that Kant's likening of testimony to experience is a *negative* point about the limits of testimony as a source of knowledge. Pasternack argues that knowledge on Kant's view requires certainty, and we can only be certain of necessary propositions known *a priori* (such as the truths of mathematics), never of contingent empirical propositions (Pasternack [10]). These issues, too, are beyond the scope of this paper. I will take Kant's claim that testimony and experience, can both yield knowledge despite their fallibility, at face value; and I will use knowledge in the contemporary everyday sense in which testimony, perception, and reasoning can all result in knowledge.

⁴Kant [6] BL p. 246.

⁵Ibid. p. 249.

Kant's idea here seems to be that if testimony is to be a source of true belief, it must be combined with an evaluation of the speaker's credibility: that is to say, the hearer needs to assess whether the speaker has the characteristics of 'competence' and 'integrity' alluded to in the *Jäsche Logic* passage, if she is to avoid 'blind belief' which leads to error.

We will return later to the precise form that this norm for hearers' uptake of speakers' testimony takes for Kant. But let us consider first what it takes for a witness to be a credible witness. In the *Blomberg Logic*, Kant lays down the following conditions:

The subject... who asserts something, and whom I am to believe, must have three attributes above all. First, capacity to obtain an experience. The subject must therefore have rational reflection. Second, it must have the capacity to preserve faithfully the experience obtained[,] or have a good memory. Third, it must be able to declare the experience obtained, to expound it, and to acquaint everyone with it. The moral character of the subject, however, is that it also have the mental constitution to represent faithfully, not otherwise, and to declare the experiences as they were obtained by him.⁶

Here Kant identifies two sets of conditions which a credible testifier must satisfy: one set of conditions is 'logical' or epistemic in nature, relating to the testifier's cognitive capacities. The other set of conditions is moral, relating to the testifier's character:

[T]he conditions under which a subject who has obtained experiences is to be believed are grounded partly in the subject's own attributes, partly in its condition. The subject's attributes are partly *logical*, since it is capable of having, of obtaining, experiences, of knowing, and of asserting. They are partly *moral*, since the subject also has the will to assert the true, or to declare the experiences as they were.⁷

We can summarise Kant's view of credibility by saying that a testifier is *credible* if and only if she meets the following two conditions:

⁶Ibid. p. 244; see also Ibid. VL p. 898.

⁷Ibid. BL p. 245.

Competence The testifier is able to obtain, preserve, and declare an experience of the event she is testifying to.

Sincerity The testifier does not intend to deceive her audience by giving false testimony.

Both of these are necessary conditions for a speaker's testimony to transmit knowledge to her audience: it is independently plausible that in order for a speaker's testimony to result in knowledge on the hearer's part, the speaker herself must know the proposition she testifies to. This is just to say that testimony, unlike e.g. perception and reasoning, and much like memory, is not a *generative* source of knowledge; testimony does not create new knowledge but merely transmits it from one epistemic agent to another.⁸ That the speaker knows what she testifies to, is secured by her meeting Competence with respect to the subject matter of her testimony. And clearly the speaker must be sincere and not lie, or otherwise the audience won't acquire (gettier-proof) justified *true* belief based on her testimony.

We now know that Kant thinks that testimony can result in knowledge of what is testified to, that testimonial knowledge rests on the reliability of testimony, and that the reliability of testimony – its being a guide to true beliefs – depends on hearers' alertness to the credibility of speakers. Only when speakers meet the logical, as well as the moral conditions of credibility, can testimony do its job as a source of knowledge for hearers.

This is a good start for an epistemology of testimony, but there are important questions left to answer. We've noted that Kant requires that hearers should make a judgement of the credibility of their interlocutors, when accepting their testimony. Uptake of testimony should not be wholly uncritical or unreflective, or it risks taking us on the 'path toward error'. This last point seems correct, but there are familiar challenges regarding how to balance the fact that in the ordinary course of things we do tend to believe others without pausing to assess their epistemic credentials, against the intuition that credulously accepting others say-so as true *without* such assessment seems like an epistem-

⁸It is not entirely uncontroversial that testimony is *never* a generative source of knowledge. Lackey argues that there are cases where a hearer can acquire testimonially based knowledge even if the speaker does not know the proposition she testifies to: her example is of a creationist teacher who teaches the (let us suppose, true) evolutionary theory, but does not believe – ergo, does not know – it (Lackey [11]). But clearly the pupils do come to know the truths of the evolutionary theory based on the teacher's tellings. If this is our view, then we can understand the role of Competence differently: it is a further *epistemic* guarantee that the proposition testified to is true, since Competence guarantees that the event is reported truthfully.

ically irresponsible, unreliable way to form beliefs. The norms of testimonial uptake in an ideal epistemology of testimony would respect both of these data: they would describe norms which don't place unrealistic epistemic demands on hearers, but which also do not permit gullibility.

In the next section we will consider in detail what Kant has to say about the norms that govern testimonial uptake, and we shall see that the principles we can attribute to Kant succeed in getting very close to this ideal, if not fully achieving it.

2 Moral and historical belief

To make progress on the question of testimony from the hearer's perspective, we should attend to a distinction that Kant draws between two different attitudes that a hearer can take towards a testifier. One attitude is epistemic, the other is a moral attitude.

Belief toward a person . . . is either moral or historical. The former consists in trusting the honesty of the other, although he has not given any statement or story. The other consists, however, in the holding-to-be-true of what the other asserts, merely because he has affirmed it . . . Moral belief, however, requires that one have a particular trust toward a person, trust that he is good and honest.⁹

'Historical belief' refers to the acceptance of a speaker's testimony as true: it is a propositional attitude to the content of a speaker's testimony. *Moral belief* refers to the belief that another person is morally upright, good, and honest. Historical and moral belief also have their contraries, *moral unbelief* and *logical unbelief* respectively:

Unbelief can be taken in the moral sense, when one lacks all trust in the moral character of a witness. He who has such moral unbelief holds all men to be liars, he trusts no man, he believes no promises, but instead always doubts the uprightness of other people. Moral unbelief is opposed to the moral foundation of all human society Incredulity [*Ungläubigkeit*], however, is distinct from unbelief [*Unglaube*]. It consists in the fact that one does not want to give approval to the logical grounds of testimony. This

⁹Kant [6] BL p. 242.

incredulity can also be called logical unbelief, and thus it is quite different from moral unbelief.¹⁰

In sum, historical belief towards a speaker consist of believing what she says, and incredulity entails refusing to believe anyone's word. Moral belief towards a speaker consists of having trust in the goodness of her character, and moral unbelief is the very opposite; moral unbelief is morally objectionable because it is 'opposed to the moral foundation of all human society'. Systematic distrust in other's honesty would be inconsistent with basic social practices like promising and contracting with others.¹¹

2.1 Moral belief and the presumption of sincerity

We can see that there is a close connection between Moral Belief and the Sincerity of speakers: if it is morally objectionable to operate on the assumption that others are insincere – this would be moral *unbelief* – then we ought to operate on the assumption that people satisfy Sincerity; that they are not out to lie and deceive us. Kant seems to be drawing just this connection when he propounds what he calls the 'Rule of fairness':

As for other things that concern the credibility and honorability of witnesses who make assertions about experiences they have obtained, everyone is taken to be honorable and upright until the opposite has been proved, namely, that he deviates from the truth. According to the well known rule of fairness: *quilibet praesumitur bonus, donec probetur contrarium*. [Everyone is presumed good until the opposite is proved.]¹²

Returning to the norms that govern testimonial uptake, and the assessments of credibility which hearers should make in trusting speakers for their word, we can now see that it is not the case that hearers need to assess speakers' Sincerity when accepting their testimony. Rather, since everyone is "presumed good until the opposite is proved," something like the following principle must govern testimonial uptake:

¹⁰Ibid. pp. 248–249.

¹¹See also Kant [6] VL p. 900.

¹²Ibid. BL p. 246.

Presumption of Sincerity You ought not to believe that someone's moral character is flawed, until you have good reason to do so.

This principle captures Kant's rule of fairness and rules out moral unbelief. It says that hearers are not under any requirement to assess their interlocutors' sincerity; as a matter of moral principle, other people's sincerity ought to be taken as a default. Note however that Presumption of Sincerity is not a principle that by itself gives hearers *epistemic licence* to accept speakers' testimony: I can wholeheartedly believe that my friend has a wonderful moral character while thinking that she is unfortunately ignorant of what she is talking about and thus, however much I trust her, I shouldn't believe what she says. That someone is sincere does not by itself give me any reason to believe that she is speaking from knowledge – that she is *competent* – and thus, it gives me no epistemic reason to take her testimony as true. So Presumption of Sincerity does not license the kind of credulity or epistemic irresponsibility that we hoped to avoid.

3 Default acceptance and Competence

We now know how hearers ought to conduct themselves when it comes to the Sincerity condition of credibility. But what about the epistemic Competence condition – are there materials in Kant for an analogous presumption of Competence? Recall that Competence claims that if a speaker's testimony is to result in knowledge on the hearer's part, then (1) we must have the capacity to obtain experiences; (2) our memory must also be faithful enough to maintain and to preserve the experiences we have had, without their being destroyed by imagination; and (3) in our language we are also capable of making others acquainted with experiences obtained.¹³

There is a temptation to think that we can quite safely assume that our interlocutors, insofar as they are normal epistemic agents, meet these conditions; Competence as it is here described sets such a low bar that it is almost trivially satisfied by any ordinary speaker. Thus, one might think, if for all we know most people whose testimony we hear are Competent, then we might as well work under the assumption that they are Competent until we are proven wrong. And if we can presume that people meet Competence, and are also obliged to assume that they meet Sincerity, then we are by default warranted

¹³Ibid. p. 245.

– maybe even obliged – to believe what others testify to. This line of thought makes it tempting to attribute to Kant something like the following principle as a norm of testimonial uptake.

Default A speaker's testimony may (or ought) to be believed, unless we have good reason to think that either (i) the speaker fails Sincerity, or (ii) the speaker fails Competence.¹⁴

A principle of this kind would be a combination of the Presumption of Sincerity and an analogous presumption of Competence. Combined with Kant's claim that Sincere and Competent speakers ought to be believed, we get the principle that *all speakers* ought to be believed, in the absence of reasons to think that the presumptions of credibility are false. There is no need for a hearer to have positive reasons based on an examination of the speaker's epistemic credentials to think that the speaker's report is likely to be knowledgeable and true, in order for her uptake to be epistemically justified.

I believe there is some truth to the thought that in certain contexts hearers may take their interlocutors' Competence for granted, and that there is nothing epistemically irresponsible about this. As we noted earlier most of our uptake of others' testimony is unreflective and immediate; to claim that this manner of uptake is inconsistent with acquiring justified belief or knowledge based on testimony would leave us with the skeptical conclusion that most of what we thought was testimonial knowledge, is in fact not so. So it must be possible at least sometimes to accept testimony unreflectively, *and* thereby gain testimonial knowledge. But this line of thought does not take us all the way to the general principle that *regardless of context*, testimony may be accepted as true by default. It only gets us to the claim that in some contexts it is fine to take things on other people's say-so, since Competence is trivially satisfied in those contexts. But if the standards for Competence *vary* from one context to another, then there may be other contexts where an assessment of the speaker's epistemic abilities is required before we may accept her testimony; where it would be wholly unreliable and irresponsible to take things on other's say-so unreflectively.

The context-dependency of our uptake of testimony is something that Kant

¹⁴Gelfert argues that Kant's commitment to his moral 'rule of fairness', or what I have called 'Presumption of Integrity', entails a default acceptance of testimony (Gelfert [4]). Section 2 showed that a default assumption of another person's moral uprightness is compatible with the belief that they don't meet Competence. So the rule of fairness is compatible with the denial of Default. Section 3 will also show that there are no grounds for Default in Kant.

himself is alert to. Consider the following comments on how the content of what is said affects our response to a speaker's testimony:

Sometimes we believe just anyone, be he who he will, merely because he says something, without attending to his person, or to his *caractère* or social position. E.g., when someone says it snowed in his area, or many ships arrived, then we simply believe this without investigating much; in these cases we do not judge about the credibility of the witnesses, then. But in certain matters, again, we do not believe just anyone, but rather we demand a particular credibility of him who puts something forward as true before we accept such a thing and hold it to be true.¹⁵

And:

There are certain assertions... regarding which one need not investigate the credibility of the witness at all, but which one rejects before any investigation. This seems at first glance to be very obstinate, but it is also certain. Cognitions that are completely and utterly and obviously false do not need any investigation at all and in fact deserve complete disdain.¹⁶

The idea behind these remarks seems to be that not every instance of testimony deserves the same reception. Sometimes we accept testimony unreflectively and “do not judge about the credibility” at all; sometimes we require very strong proof of the speaker's credentials before we are willing to accept their testimony as true.

Now the suggestion under consideration was that since Competence as Kant initially defines it is so easy to come by, we can presume that speakers are Competent as well as Sincere, and so may follow a principle like Default when accepting the testimony of others. In the first kind of case mentioned in these passages, where the experience reported is of a very ordinary sort, it is true that Competence comes as a default; there seems to be no great risk of error in presuming it. In these cases, unreflective default acceptance of testimony seems permissible – not because the need to make a judgement about the speaker's credibility is waived but because Competence, given the subject matter, is trivially satisfied by ordinary speakers. But when the subject matter of the testimony is more controversial or complex – e.g. what is asserted seems to contra-

¹⁵Kant [6] BL p. 242.

¹⁶Ibid. p. 247.

dict common sense, or knowing the thing requires expertise – the requirement to be more reflective and discriminating in one’s acceptance of testimony kicks in. If our uptake of testimony is to be a reliable means to true beliefs and to knowledge, we must make a conscious judgement whether the speaker has Competence in this domain.

The upshot is that the presumption of Competence is a reasonable epistemic principle to abide by in a limited range of cases; to find the limits of this range we ought to look to which cognitive abilities are required for Competence in the relevant domain, and whether it is reasonable to presume without further investigation that the speaker has these abilities.¹⁷ So instead of Default, we should take Kant to be advocating a principle along the following lines.

Assessment of Competence You should not accept another person’s testimony unless you can reasonably be confident that she is Competent and knowledgeable about the subject matter of her testimony.

This allows for context-sensitivity in Competence itself, and in how extensive an examination into the speaker’s epistemic credentials is necessary for acceptance of testimony to be appropriate: When a long-time friend tells you what she had for breakfast this morning, you can reasonably be confident that she knows what she is talking about without much of an inquiry into her accuracy about the subject matter. But when a stranger on the train tells you that she met Elvis in the casino last night, you would be well advised to consider carefully whether this person’s memory, or senses, or capacity to report her experiences truthfully are in normal working order.

4 Testimony and epistemic progress

This combination of a Presumption of Sincerity and an Assessment of Competence seems to me like an eminently reasonable beginning for a fuller account of the epistemology of testimony. It captures Kant’s condemnation of a

¹⁷Fricker (1995) suggests that we should “disaggregate” our uptake of testimony in just this way: “Intuitively, some people or types of people on some topics are reliable, others on others aren’t. Thus the key to the epistemology of testimony is: disaggregate... both regarding the question whether and when we may rightly trust without evidence, and regarding the empirical confirmation of speakers’ trustworthiness.” Fricker’s own view on this is similar to how I’ve described Kant’s: “There is a default presumption in favour of competence, but only with respect to a restricted range of subject matters: those where what we all know about people and their normal capabilities shows that competence may indeed be presumed.” (Fricker [3] p. 407)

distrustful attitude towards others' moral character, and it also captures his requirement for epistemic responsibility in our acceptance of other's testimony. Before concluding, let us return to the First Critique where Kant describes the testimony of others as a 'touchstone for truth', as a corrective for the validity of our own judgements. Here is the passage in full.

Truth... rests upon agreement with the object, with regard to which, consequently, the judgements of every understanding must agree... The touchstone of whether taking something to be true is conviction or mere persuasion is therefore, externally, the possibility of communicating it and finding it to be valid for the reason of every human being to take it to be true; for in that case there is at least a presumption that the ground of the agreement of all judgements, regardless of the difference among the subjects, rests on the common ground, namely the object, with which they therefore all agree and through which the truth of the judgement is proved.¹⁸

By 'persuasion' Kant refers to a situation where an agent mistakenly thinks that she has good reasons for taking some proposition to be true, when in fact she does not. The subject's mistake in persuasion is due to something which is idiosyncratic to her, which makes the reasons appear to her to be very good reasons indeed; for example her false belief about the weight of her reasons. Thus the subject's grounds for her belief have merely 'private validity', not universal validity for all rational beings. Their mere private validity can be revealed by other people's disagreement: if others are not convinced by your reasons for taking some proposition to be true, then other things being equal, you have reason to think that there is something wrong with your judgement about the weight of those reasons. Others' disagreement gives you a reason to think that your belief is false. Other people's judgement is of course not an infallible check on our own beliefs since the others might be mistaken; similarly, the lack of disagreement is no guarantee that what we think is true since we might all be mistaken. But it is, according to Kant, at least a helpful negative criterion that should help us detect and avoid falsehoods, insofar as we think that other people's judgement is no worse than ours.¹⁹

¹⁸ Kant [7] A 812/B 849.

¹⁹ To think otherwise would of course be incredulity, or logical unbelief: "Logical egoism is a *selfish prejudice*... which takes as dispensable the criterion of truth, *to compare one's opinions with those of other men*." (Kant [6] DWL p. 740)

5 Conclusion

Let us take stock and conclude. We've argued that Kant's epistemology of testimony involves at least the following two norms which guide hearers in their acceptance of testimony.

Presumption of Sincerity You ought not to believe that someone's moral character is flawed, until you have good reason to do so.

Assessment of Competence You should not accept another person's testimony unless you can reasonably be confident that she is Competent and knowledgeable about the subject matter of her testimony.

The former of these is a general moral principle, not exclusively tied to testimonial exchange. But it is obviously relevant to testimony, since we couldn't acquire knowledge from speakers if we did not trust them to be speaking truthfully. The latter is an epistemic principle directly relevant only to testimony: Assessment of Competence denies that we have epistemic licence to accept others' testimony without an eye to whether they are likely to be knowledgeable about the subject matter of their testimony. Following this norm ensures that our acceptance of testimony will be a reliable means to true beliefs, not to "blind belief" and error. We've also argued that these are just the principles for testimonial uptake that hearers need to follow, if testimony is to serve its role as a means to epistemic improvement. So in addition to having strong textual support, these principles pass the litmus test of being consistent with – and indeed, supported by – Kant's views on the role of testimonial practice in our epistemic lives. They also respect the fact that sometimes our uptake of testimony is immediate and unreflective, but that there are contexts where it would be epistemically irresponsible and gullible to simply accept what others say without a judgement of their epistemic credentials; thus meeting the desiderata of not imposing unreasonable epistemic demands where they seem out of place, while also not licensing gullibility.

References

- [1] Faulkner, P. (2014) 'A Virtue Theory of Testimony', in *Proceedings of the Aristotelian Society* 114/2: 189–211.
- [2] — (2014) 'The moral obligations of trust', in *Philosophical Explorations* 17/3: 332–345.

- [3] Fricker, E. (1995) ‘Telling and trusting: Reductionism and Anti-Reductionism in the Epistemology of Testimony’, in *Mind* 104: 393–411.
- [4] Gelfert, A. (2006) ‘Kant on Testimony’, *British Journal for the History of Philosophy*, 14: 627–652.
- [5] — (2010) ‘Kant and the Enlightenment’s Contribution to Social Epistemology’, in *Episteme*: 79–99.
- [6] Kant, I. (1992) *Lectures on Logic*, J. Michael Young (trans.). Cambridge: Cambridge University Press.
- [7] — (1998) *Critique of Pure Reason*, P. Guyer & A.W. Wood (trans.). Cambridge: Cambridge University Press.
- [8] Schmitt, F.F. (1987) ‘Justification, Sociality, and Autonomy’, in *Synthese* 73: 43–85.
- [9] Shieber, J. (2010) ‘Between Autonomy and Authority: Kant on the Epistemic Status of Testimony’, in *Philosophy and Phenomenological Research*, 80/2: 327–348.
- [10] Pasternak, L. (2014) ‘Kant on Opinion: Assent, Hypothesis, and the Norms of General Applied Logic’, *Kant-Studien* 105/1: 41–82.
- [11] Lackey, J. (2006) *The Epistemology of Testimony*, E. Sosa (ed.). Oxford University Press.
- [12] — (2006) ‘It Takes Two to Tango: Beyond Reductionism and Non-Reductionism in the Epistemology of Testimony’, in *The Epistemology of Testimony* J. Lackey & E. Sosa (eds.). Oxford University Press.

Solipsism in the *Tractatus**

Sam Davies Udina

University of Oxford

I am a solipsist, according to the standard usage of the word, if I believe that the world is nothing but an object of my experience.

Wittgenstein places his remarks on solipsism in the *Tractatus* at the 5.6s. He says: “In fact what solipsism *means*, is quite correct, only it cannot be *said*, but it shows itself.”¹ We must perform three interpretative tasks: determine whether what Wittgenstein takes solipsism to mean is what we usually take solipsism to mean; make clear why Wittgenstein takes it to be correct; and explain why Wittgenstein thinks it is one of the things that can be shown but not said.

It is first important to characterise the Russellian views that are implicitly being attacked in the remarks on solipsism. For Russell (or rather, for the relevant stage of Russell’s thought), I can name only that with which I am acquainted, and I am acquainted only with that which I can name. More generally, I must be acquainted with the components of any proposition which I judge. I am acquainted with my sense-data, for instance, so I can judge propositions the components of which are my sense-data. I am also acquainted with the logical objects corresponding to the universal and existential quantifiers, expressed in language by the logical constants “all” and “some”. This latter assumption allows Russell to hold that I can talk about objects with which I am not acquainted by means of definite descriptions: for definite descriptions use only quantified propositions and may not involve any names. Hence although I am not acquainted, for instance, with A’s private experience, and hence I cannot name it, I can talk about it by means of an expression of the sort ‘the unique experience being had by A’. This use of bounded variables thus plays a crucial part in acquiring knowledge by description, and it allows me to escape the ‘prison’ of my present experience. For Russell, the limits of what is knowable lie outside the realm of my experience, and thus solipsism is wrong. (Diamond calls Russell’s a ‘two-limit’ realism: the limits of my experience are different

*Delivered at the BJUPS Winter Conference, 14–15 January 2017 at UCL.

¹Wittgenstein [6] 5.62.

from the limits of the world.)²

If the notions of ‘all’ and ‘some’ are primitive, as they are for Russell, then it is left somewhat unexplained why, for instance, ‘ $f(a)$ ’ follows from ‘for all x , $f(x)$ ’, or why ‘there is x such that $f(x)$ ’ follows from ‘ $f(a)$ ’. Indeed these entailments still hold even if ‘ a ’ names an object with which I am not acquainted. But that is a case in which I don’t even *understand* the sentence ‘ $f(a)$ ’, since it involves a name for an object with which I am not acquainted. Thus it is possible, on this view, to have sentences I do *not* understand entail sentences which I *do* understand.

This is unacceptable for Wittgenstein. For him, if a sentence I understand follows from another then they are both in logical space – that is, they are in the space of sentences constructed from the language I understand – and hence it is incoherent to suppose that I understand one and not the other. No; Wittgenstein’s notion of a quantifier works in a very different way from that which has been discussed above. The logical constants ‘all’ and ‘some’ do not refer to logical objects: this is, after all, his *Grundgedanke*.³ Rather, they must be seen as a construction from singular sentences (i.e. sentences of the form ‘ $f(a)$ ’). In fact, the construction is straightforward (at least for the finite case): ‘all’ will simply involve the conjunction of singular sentences, and ‘some’ the disjunction of singular sentences. This dissolves the mystery around the question why ‘ $f(a)$ ’ follows from ‘for all x , $f(x)$ ’ and why ‘there is x such that $f(x)$ ’ follows from ‘ $f(a)$ ’.

Thus, for Wittgenstein, the introduction of variables is nothing but notational convenience; and, crucially, quantifiers are not magical fishing nets which allow me to range over objects which I cannot name. In particular, I cannot use quantifiers to talk about A ’s private experience, in the way Russell expected I could. This is why Diamond talks about the private language argument being, to an extent, in the *Tractatus*: my own beetle hasn’t been shown to be out of the language game yet, but other people’s beetles have been. A more general consequence of Wittgenstein’s position regarding quantifiers, however, is that Russell’s attempt to escape solipsism by use of quantifiers fails, since they do not allow me to have knowledge by description, which was my key out of the prison.

It is not obvious that Wittgenstein’s theory of quantifiers is correct. If a univer-

²Diamond [2].

³Wittgenstein [6] 4.0312.

sal quantifier is analysed at some level as a conjunction of singular sentences, then at the very least it must follow from the conjunction of the same singular sentences. But ordinarily we do not take the set containing ' $f(a)$ ', ' $f(b)$ ', ' $f(c)$ ', and so on to entail 'for all x , $f(x)$ ' unless provided with a further sentence saying that there is nothing in the domain but a , b , c , and so on. Within the framework of the *Tractatus*, however, this is not a weakness. For Wittgenstein the above further sentence would be nonsensical: we cannot in logic say "the world has this in it, and this, but not that".⁴ Another worry for the theory might be whether it can be extended to quantification over infinite domains, since infinite conjunctions and infinite disjunctions are in general not allowed. But there seems to be no objection in principle to an extension to the infinite case: all that is needed is to analyse 'for all x , $f(x)$ ' as being true whenever all the relevant singular sentences are true, and 'there exists x such that $f(x)$ ' as being true whenever *at least one* relevant singular sentence is true.

The above, of course, does nothing to establish the truth of solipsism. To see how Wittgenstein does this, it must first be understood what, for Wittgenstein, solipsism 'means' ('*meint*'). Here different philosophers suggest different interpretations. I favour a straightforward interpretation: that for Wittgenstein, what solipsism means is that the world is *my* world. After all, he explains *why* the world is my world immediately after saying that what solipsism means is quite correct, suggesting that both ideas are closely connected.⁵ This characterisation of solipsism is consistent with the notion employed by other philosophers, and consistent, if not synonymous, with the definition that started this essay. In any case, precisely what 'the world is *my* world' means must still be explained; but if this interpretation is right, then it is enough, in order to establish the truth of solipsism, to find an argument that has 'the world is *my* world' as a conclusion.

Such argument is suggested by Anscombe: "the limits of my language mean the limits of my world; but all languages have one and the same logic, and its limits are those of the world; therefore the limits of my world and of the world are one and the same; therefore the world is my world."⁶

The premises to this argument certainly all sound very Wittgensteinian. They will be examined soon. First, though: the discussion conducted above concerning Wittgenstein and Russell's respective accounts of quantifiers adhered to a

⁴Ibid. 5.61.

⁵Ibid. 5.62.

⁶Anscombe [1].

premise I made explicit at the start, that I can name something just in case I am acquainted with it. The trouble with adhering to this is that, in combination with the conclusion of the argument just presented, it leads to a kind of solipsism that simply isn't very interesting. In the literature it is usually labelled 'empirical solipsism', and it was almost certainly not the solipsism to which Wittgenstein was ascribing correctness. For, according to most philosophical uses of the word 'acquainted' (including Russell's), there is plenty with which I am not acquainted. It is not only Bismarck's private experience that I am not acquainted with, but also Bismarck himself. (I can talk about Bismarck, according to Russell, only because 'Bismarck' is in fact just standing for a definite description.) If my lack of acquaintance with Bismarck entails I cannot name him, and if the world is my world so that there is nothing that my language cannot name, then Bismarck is not part of the world. And likewise for the table on which I'm resting my forearms. This is silly, and it is not Wittgenstein's view in the *Tractatus*.

The reason it is not is that Wittgenstein abandoned the aforementioned premiss before writing the *Tractatus*, most likely while in Norway. Thus the Tractarian objects, which are the prototypes of bearers of proper names, need not be objects of my acquaintance – which is just as well, because wherever these aspect-less objects are to be found (if at all, says Ishiguro),⁷ it certainly doesn't seem to me as if I've ever experienced them. In an unpublished lecture, Potter gives an example: we may construct a system of coordinates by which all the points in a given space (which are Tractarian objects) are given a name, including those with which I am not acquainted. Thus even if the world is my world, it might include things with which I am not acquainted.

Back to Anscombe's reconstruction of the argument for solipsism. The first premise is that the limits of the language mean the limits of my world. This is 5.6, the central proposition on which all the remarks on solipsism are comments. Certainly if my world includes only those things which I can name, then its limits coincide with the limits of my language. Let us jump to the third premise, that the limits of logic are the limits of the world. Using the word 'logic' as Wittgenstein does, this seems true: certainly nothing in the world can be beyond logic, and to say that there is something in logic that is not in the world must surely be nonsense.

Now; the second premise is that all languages have the same logic. Precisely what it means to say that a language 'has a logic' is unclear. If the premise is

⁷Ishiguro [3].

saying that all languages are logical, i.e. share the logical form that the world has, then this seems to be true, at least within the Tractarian framework. But then it is not obvious at all how this premise slots in with the other two to reach the conclusion. Alternatively, the premise might be saying that all languages are such that, for any object, a name that names that object can be introduced, and in that sense share a logic the limits of which are the limits of the world; then the premises do seem to slot together, but it becomes doubtful whether the second premise is true. In what sense does our language have the potential to name every object? Propositions are truth-functions of elementary propositions, each of which asserts the existence of an atomic fact, which is a combination of objects. So every proposition includes, at some level of analysis, names for Tractarian objects. But this gives no guarantee that our language includes, or could include upon expansion, names for all objects. In other words, what Proposition 6 asserts is that any proposition of our language is a truth-functional construction from elementary propositions; but there is no guarantee that the converse to Proposition 6 holds, namely that any truth-functional construction from elementary propositions that conforms to the general form of the proposition is a proposition of our language. I think the justification for the converse of Proposition 6, which in turn is the justification for claiming that our language is such that indeed it can have names for all objects, comes straight from Wittgenstein's central assumption: that the world is represented by our language. Any state of affairs – involving any object – can be described by a proposition of our language, and hence our language has the potential to name every object. The second premise, and hence the conclusion of the argument for solipsism, is as true as Wittgenstein's central assumption is.

How are we to understand this solipsism? Not, I have said, as the empirical solipsism against which Descartes and Russell battle. The remarks on the metaphysical self provide some help. Wittgenstein says: "There is no such thing as the subject that thinks or entertains ideas"; and also: "there is therefore really a sense in which in philosophy we can talk of a non-psychological I."⁸ The philosophical I is not the man, not the human body or the human soul of which psychology treats, but the metaphysical subject, the limit of the world—not a part of the world."⁹

There is a distinction between two notions, neither of which corresponds to

⁸Wittgenstein [6] 5.361.

⁹Ibid. 5.641.

anything in the world, but one of which does exist as the limit of the world – not that ‘exists’ is the right word at all. First, there is a rather clear argument readily available as to why the ‘thinking, presenting’ subject does not exist. If it did, then it would be able to infer its own existence as a necessary consequence of any thought it had, and this sort of characterisation violates Wittgenstein’s premiss that the only necessity is logical necessity.

The metaphysical self, on the other hand, is the limit of the world: any attempt to talk about it will be nonsensical (including this one). I follow Williams in thinking that Wittgenstein says that philosophy can talk of a non-psychological I only in the sense in which philosophy, according to the *Tractatus*, can talk about anything: nonsensically.¹⁰ Having said that, the most fruitful way of thinking about the metaphysical self I have found is in comparison with the eye-and-field-of-sight analogy.¹¹ We are told that the eye is not part of the field of sight. We are also told that the I of solipsism – i.e. the metaphysical self – shrinks to an extensionless point, leaving reality coordinated with it. I think there is a way of understanding the field of sight as coordinated with the eye, and the eye, though not part of the field of sight, as being at its limit. And I think we can interpret the remark on the extensionless point analogously: the same happens with reality and the self. In fact, the analogy holds even further. If we could gain a ‘sideways’ view, we could see both the eye and the field of sight. Of course this is not only impossible, it is nonsensical: we cannot ‘see’ a field of sight from anywhere other than its limit, the eye. And it is nonsensical in a similar way as it is to suppose that we can gain a ‘sideways’ view and see the world and the self as its limit. This would require stepping out of the world, which is nonsensical. These are one and the same impossibility.

We are now in a position to understand why Wittgenstein says solipsism is correct. When considering Russell above, I spoke about his picture of the world having two limits, one being the limit of my experience and the other being the limit of the world. The Tractarian solipsism is not shrinking the second limit to make it coincide with the first limit – it is expanding the first limit to make it coincide with the second: there is nothing which my language cannot name. This is why solipsism “strictly carried out” coincides with pure realism.¹² And because the world is representable, because it can be projected onto language, there is the metaphysical subject, which corresponds to the ‘my’ of ‘my lan-

¹⁰Williams [5].

¹¹Wittgenstein [6] 5.6331.

¹²Ibid. 5.64.

guage', and which, as McGinn puts it, is in the position to say, 'I think...'.¹³

As was noted above, Diamond suggests that the remarks on solipsism amount to an embryonic form of a private language argument. She writes, "if one identifies that insight (the rejection of Russell's conception of access to what lies 'beyond' experience) as what solipsism really means to say, then we can say, as Wittgenstein does, that what solipsism means to say is correct. (This is not to say that solipsism is correct.)"¹⁴ Though, as we saw above, it is certainly true that Wittgenstein's account of quantifiers leads to this partial private language argument, I think that to read the remarks on solipsism as simply expressing this insight understates the extent to which solipsism is embraced in the *Tractatus*. Solipsism is correct not only in saying that it is incoherent to think that there is a realm of private experience containing objects which I cannot name, as Diamond would have it, but in saying that it is incoherent to think that there are *any objects at all* which I cannot name. Russell need not worry about escaping the limits of his experience: there is only one limit, and there is nothing beyond it, but that is because it already includes everything.

The last interpretative task is to explain why Wittgenstein says that although what solipsism means is correct, it cannot be said, but rather shows itself. But this is straightforward. It must clearly be nonsense to say that I can name everything there is in the world. For one thing, talking about 'everything there is in the world' is already nonsensical, since we cannot in logic say 'the world has this in it, and this, but not that'.¹⁵ But there is an even more direct route to the conclusion that what solipsism means cannot be said. Once it has been understood that the world is my world, it must surely be recognised as nonsensical to say so: what else could the world be, but my world?

References

- [1] Anscombe, G.E.M. (1971) *An Introduction to Wittgenstein's Tractatus*. London: Hutchinson & Co.

¹³McGinn [4].

¹⁴Diamond [2].

¹⁵Wittgenstein [6] 5.61.

- [2] Diamond, C. (2000) ‘Does Bismarck have a beetle in his box? The private language argument in the Tractatus’, in A. Crary & R. Read (eds.), *The New Wittgenstein*. London: Routledge.
- [3] Ishiguro, H. (1969) ‘Use and Reference of Names’, in P. Winch (ed.), *Studies in the Philosophy of Wittgenstein*. London: Routledge & Kegan Paul.
- [4] McGinn, M. (2006) *Elucidating the Tractatus*. Oxford: Clarendon Press.
- [5] Williams, B. (1973) ‘Wittgenstein and Idealism’, in *Royal Institute of Philosophy Lectures 7*: 76–95.
- [6] Wittgenstein, L. (1971) *Tractatus Logico-Philosophicus*, D. F. Pears & B. McGuinness (trans.). London: Routledge.

A Host of Uncontroversial A Priori Knowable Contingent Truths*

Oliver Hurcum

University of Oxford

1 Introduction

A priori knowledge is knowledge resting on *a priori* justification. A belief in some proposition p is justified *a priori* iff (if and only if) it is reasonable to believe p solely on the basis of understanding p ; ‘All bachelors are unmarried’ serves as a paradigm example. A priori knowledge is contrasted with *a posteriori* knowledge: knowledge resting on *a posteriori* justification. A belief in some proposition p is justified *a posteriori* iff it is not justified *a priori*: iff some kind of experience beyond understanding P is necessary to make the belief reasonable. ‘Water is H_2O ’ is known *a posteriori*, as are propositions such as ‘I am in pain?’ or ‘I am happy’. A truth is contingent iff it could have been false: iff it is not necessarily true.

There is considerable debate among epistemologists over whether there is such a thing as an *a priori* knowable contingent truth. *Prima facie*, it appears we should answer negatively; if merely understanding the proposition is sufficient to know its truth-value then it surely has this truth-value necessarily rather than contingently. However, Kripke claimed to have found an example, and since then numerous others have offered up their own candidates for the title of the contingent *a priori*.¹ Unfortunately, all these putative cases have proven highly controversial and the dispute is widely considered to be far from settled.

I now believe, somewhat tentatively, to have found not one but a host of uncontroversial *a priori* knowable contingent truths; that the external world exists, that other minds exist, that I am the same person I was a moment ago, and so on. Denials of sceptical hypotheses, indeed, are a rich source of such truths. If I am right, the debate may finally be brought to a close.

The essay is split into three parts. First, I consider Kripke’s famous original

*Delivered at the BJUPS Winter Conference, 14–15 January 2017 at UCL.

¹Kripke [11].

example and explain why it is thought to fail.² Then, I attempt to demonstrate that my suggestions are indeed uncontroversially both contingent and *a priori*. My proof that they are *a priori* is, however, what might be called non-constructive; I show, if my proof is successful, that they are undoubtedly known *a priori* without explaining how it is that they are so known. I concede that I am unable to offer such an explanation, but the third part of this essay offers some thoughts on how we might go about finding one.

2 Kripke

It is Kripke who first broke with tradition and suggested that *a priori* knowable contingent truths were possible.³ He points out that apriority is an epistemological notion while necessity is a metaphysical notion, and so “it’s certainly a philosophical thesis, and not a matter of obvious definitional equivalence, either that everything *a priori* is necessary or that everything necessary is *a priori*” since the two notions “are dealing with two different domains”.⁴

Attempting to demonstrate that the thesis is false, he invites us to consider the case where we introduce the term ‘one metre’ as the name of that which happens to satisfy the description ‘the length of the stick S at time t’. Crucially, we do not intend ‘one metre’ to mean or abbreviate ‘the length of S at t’; the two are not definitional equivalents. Rather, the description is used merely to fix the reference; we use the description to pick something out and it is that thing to which the name is stipulated to refer. We are then presented with the proposition ‘S at t is one metre long’. Given our stipulation, this seems straightforwardly to be known *a priori*. Moreover, the proposition also appears contingent, for there are possible worlds where S at t has a different length, a length different to the one we have named ‘one metre’.

The most influential critique of Kripke comes from Donnellan.⁵ Donnellan does not so much dispute Kripke’s claim to have found an *a priori* contingent truth, but claims that the truth Kripke has found is philosophically uninteresting. It is philosophically uninteresting because it says nothing substantive about the world; it is a purely linguistic truth, and presumably believers in the

²Ibid.

³Ibid.

⁴Ibid.

⁵Donnellan [6].

contingent *a priori* want something more than this. His rebuttal is premised on there being a distinction between knowing “that certain sentences express truths” and knowing “the truth of what they express.” He gives the example “The oblateness of Mars is .003”.⁶ By reading this sentence in an authoritative article, I can come to know that this sentence is true. If, however, I don’t know what ‘oblateness’ means then I certainly don’t come to know the truth that is expressed. Donnellan then claims that merely knowing that a sentence expresses a truth, without knowing the truth that it expresses, is not to genuinely know anything about the world since most will agree that I do not actually know anything about Mars here. Foreign languages provide a host of other examples.⁷ Knowing that “*Le Mineur est une piè ce russe célèbre que l’auteur Denis Fonvazine a écrite au dix-huitième siècle*” expresses a contingent truth, because I am reliably told so by a fluent French speaker, and without understanding the French, is quite clearly not to actually know any non-linguistic fact.

It is Donnellan’s contention that “S at t is one metre long” is like the French above; we come to know that it is true without actually coming to know anything substantive.⁸ This is because we do not know, in Kripke’s example, what ‘one metre’ means. It was crucial to Kripke’s argument that the expression does not mean the description that was used to fix its referent; rather, it just means its referent. But we cannot know what this referent is *a priori*, without any experience, only *a posteriori*, by actually being acquainted with it. Suppose, in order to become acquainted with it, that we went and measured the stick in some available unit (inches, hand-spans, etc.) and found that it is, say, five hand-spans long. We then know that “one metre” means ‘five hand-spans’. But the proposition ‘S at t is five hand-spans long’, which is just what ‘S at t is one metre long’ means, is known *a posteriori*; we do not know it *a priori* and it is therefore not an *a priori* knowable contingent truth.

⁶Donnellan [6].

⁷Ibid.

⁸Ibid.

3 A Host of Uncontroversial A Priori Knowable Contingent Truths

Following Kripke's lead, Plantinga, Kaplan and Evans famously touted their own candidates. Plantinga offers the proposition 'I exist',⁹ Kaplan suggests 'I am here now',¹⁰ and Evans recommends all propositions of the form 'If actually φ , then φ '.¹¹ Both Plantinga's and Kaplan's examples may be challenged on the grounds that they are in fact justified on the basis of at least introspective experience and are therefore known *a posteriori*,¹² while Evans' proposal may be criticised for its reliance on the implausible idea that 'actually' is a device for rigidifying reference rather than an indexical.¹³ I am not committed to the view that these criticisms are successful, but they have certainly made the candidates listed here deeply controversial. My contention is that there exists a host of uncontroversial examples around which proponents of the contingent *a priori* can more effectively rally. I suggest the following.

3.1 The external world exists

Unless we are sceptics, we accept that we may count 'the external world exists' among the known truths. Moreover, I do not believe that any of us are actually sceptics. We do know that the external world exists; that it is not an illusion constructed by an evil demon or super-scientist. If we are being honest, we interpret the sceptical challenge not so much as threatening our claim to know such a thing but as revealing that we don't have a perfect understanding of what knowledge is. We do know that the external world exists, that much is a primitive fact, but we are embarrassed that no analysis of 'knowledge' seems to include this proposition among what is known. It is not our knowledge which is being called into question, but our analyses of 'knowledge'.

I therefore start from the assumption that 'the external world exists' is a known truth. I also take for granted that to know p , a necessary (though by no means sufficient) condition is to have a justified belief that p . The question is then

⁹Plantinga [12].

¹⁰Kaplan

¹¹Plantinga [12], Kaplan [10], Evans [7].

¹²Fitch [8].

¹³Bostock [2].

‘What justifies our belief that the external world exists?’ Obviously, our justification cannot be *a posteriori*, for our experiences are unable to distinguish between the case where the world is an illusion and the case where the world is real. That is, no observation can be made that would justify our claim to knowledge since the sceptic can always retort ‘but how do you know that this observation was not also part of the illusion?’ If we are not justified *a posteriori*, and yet we may claim knowledge, it follows that our knowledge must be justified *a priori*.

Yet it is also obviously the case that ‘the external world exists’ is a contingent truth, for Descartes’ malignant demon or Putnam’s brains in vats are logical, though outlandish, possibilities.¹⁴ This being the case, we have here an *a priori* knowable contingent truth.

3.2 Other minds exist

‘The external world exists’ does not uniquely hold this status. Sceptical hypotheses provide a host of other examples. Consider, for example, scepticism about other minds. Unless we are sceptics, and, as I explained above, I do not believe that any of us actually are, then we accept that ‘other minds exist’ is true. Very few of us take the possibility of Chalmers’ zombies seriously; the possibility of individuals exactly like us in all physical and behavioural respects but without conscious experiences.¹⁵ Very few of us accept, for instance, that we do not know that our loved ones are conscious in the same way that we ourselves are. The spectre of scepticism, again, serves only to remind us that we do not have a satisfactory analysis of knowledge since no analysis places ‘other minds exists’ among what is known.

What justifies our belief that other minds exist? Again, our justification cannot be *a posteriori*; zombies behave in every way as though they were conscious and so no observation, no interaction with another person, would be sufficient justification for ruling out their zombie-hood. Since we agree we are justified in claiming such knowledge, our justification must therefore be *a priori*. Moreover, ‘other minds exist’ is quite clearly contingent; the existence of zombies is no logical impossibility. Here, then, is another *a priori* knowable contingent truth.

¹⁴Descartes [5], Putnam [14].

¹⁵Chalmers [3]

And there are others, all derivable from a consideration of some kind of scepticism; ‘the past exists’, ‘I am the same person I was a moment ago’, ‘other people are the same people they were a moment ago’ and so on. In each case we simply note that, because we are not sceptics, the proposition is known, is also contingent, but cannot possibly be justified *a posteriori*. *A priori* knowable contingent truths are thus not such a rare breed after all.

4 An Account of the *A Priori* Justification

An interesting question, of course, is ‘How are our beliefs, our denials of sceptical hypotheses, justified *a priori*?’ I have demonstrated that they must be so justified but I am, to my frustration, unable to explain how; I believe that this must be the subject of future research. However, I have a few considerations to offer.

4.1 Appealing to Intuitions

Here is an avenue which, though perhaps tempting, I do not believe to be fruitful. We have powerful intuitions that the external world exists, that other minds exist, that I am the same person I was a moment ago and so on. Bealer makes a robust case in favour of counting intuitions as evidence and, as he points out, using intuitions evidentially is in any case part of our standard justificatory procedure.¹⁶ To use his own example, our evidence that the traditional tripartite theory of knowledge as ‘justified true belief’ is false is our intuition that Gettier cases could occur and that in such cases we do not have knowledge.¹⁷ Furthermore, he counts intuitions as “data of reason” and not “data of experience”;¹⁸ knowledge justified on the basis of intuitions, according to Bealer, is *a priori* knowledge. Thus, one possible account of how we know *a priori* that the external world exists is that our strong intuitions constitute sufficient *a priori* justification to know such a thing.

Why is this line of argument unsuccessful? I accept that intuitions have some evidential force, but I have two concerns: one major, the other a side-concern. The side-concern is with classing intuitions as a form of *a priori* rather than a

¹⁶Bealer [1].

¹⁷Ibid.

¹⁸Ibid.

posteriori evidence. As Bealer says, “for you to have an intuition that A is just for it to seem to you that A”.¹⁹ As far as I understand it, though, this ‘seeming’ is an experience rather than a mode of understanding. I *feel* as though q is true and it is my *experience* of this feeling that I use as evidence. Our knowledge that the external world exists, if our intuitions constitute sufficient justification, would therefore be *a posteriori* on this account. My major concern, however, is that our intuitions do not have enough evidential force to constitute a justification for our claim to knowledge. Intuitions are only useful when we have a reason to believe they are reliable; my intuition that my friend was not guilty constitutes good evidence because I know my friend well, whereas my intuition that a random stranger is not guilty is no evidence at all. Crucially, there is no reason to suppose that my intuition that the external world exists, or that other minds exist and so on is any kind of reliable guide.

4.2 If there is something that I know, then I know that that external world exists

Here, on the other hand, is a line of inquiry which I feel may be helpful. My contention is that the proposition above is an *a priori* knowable contingent truth. It is knowable *a priori* because it is sufficient merely to reflect on what we mean by ‘know’ to ascertain that it is true; understanding the proposition alone is sufficient to come to know it. The proposition is nevertheless contingent because it is itself a contingent fact that the meaning of ‘know’ guarantees its truth. How this is possible will be explained. As is nodoubt already clear, this candidate for the title of the contingent *a priori* is much more controversial than those considered above, depending as it does upon an ‘alternative’ theory of knowledge. However, if we are prepared to grant it this title, it will perhaps pave the way for an explanation of the a priority of the less contentious ‘the external world exists’ for it seems, at least *prima facie*, that the latter should be deducible from the former.

What it is that we mean by ‘know’? Craig begins his book *Knowledge and the State of Nature: An Essay in Conceptual Synthesis* with the following.

The standard approach to questions about the concept of knowledge has for some time consisted in attempts to analyse the everyday meaning of the word ‘know’ and its cognates. Such attempts have usually taken the form of a search for necessary and sufficient

¹⁹Ibid.

conditions which, when measured against our reactions to examples both real and imaginary, match our intuitive ascriptions and withholdings of the title of knowledge. We are to provide, if you like, an explicit intension to fit the intuitive extension.²⁰

This standard approach, as old as Plato (circa 369 BC) and pursued with modern rigour since Gettier, has, as he points out, been unsuccessful.²¹ I have already mentioned one reason: we are compelled to place ‘the external world exists’ within the extension of the concept but can give no explicit intension that concurs with this placement. It is perhaps therefore worthwhile pursuing a different strategy. The approach Craig recommends is thoroughly naturalistic.

Instead of beginning with ordinary usage, we begin with an ordinary situation. We take some *prima facie* plausible hypothesis about what the concept of knowledge does for us, what its role in our life might be, and then ask what a concept having that role would be like, what conditions would govern its application.²²

What purpose might the concept of knowledge serve? Craig points out that early on we needed a means of flagging good informants: individuals who will reliably provide us with true beliefs that will enable us to act successfully. My survival depends on, for example, having a true belief about where my predators are and which foods I may safely eat. We are not interested, however, in flagging individuals who, reliably or not, provide us with true beliefs that can have no role in our decision making. He suggests that a good informant for me should (1) be accessible to me here and now, (2) be recognisable by me as someone likely to be right about *p*, (3) be as likely to be right about *p* as my concerns require and (4) have open channels of communication with me.²³

However, because we live in communities, I am also interested in flagging informants which are good for others and in others’ flagging informants which are good for me. We therefore begin to ‘objectivise’ what makes a good informant. This means playing down (1), (2), (4) and playing up (3); since we don’t know what level of likelihood of correctness other people’s concerns will require, we set the bar very high so that we don’t point them toward unsuitable informants with disastrous results. Craig concludes that it is this objectivised

²⁰Craig [4].

²¹Plato [13], Gettier [9].

²²Ibid.

²³Ibid. (4) has been slightly reworded.

notion of the good informant which is captured by the term ‘knowledge’.²⁴

Just how high do we set the bar? Well, as obvious as it seems, it is worth noting that we do sometimes want to flag individuals; we have not developed the concept just to never use it. The bar must therefore also be set low enough to allow that there is some knowledge. That is, there must be some propositions of the form ‘ X knows p ’, where p provides useful information that enables successful action, which are true.²⁵ Importantly, for p to provide useful information, for this information to help me to act successfully, p must say something about the external world. But p , because it describes the external world, entails that the external world exists. Assuming the principle of closure under known entailment, this leads us, *a priori*, to our desired conclusion; that if there is something that I know, then I know that the external world exists.

The worry is that if the meaning of ‘know’ guarantees that ‘If there something that I know, then I know that the external world exists’ is true, then this *a priori* proposition cannot be contingent. I suggested that, to the contrary, it is contingent because the guarantee is itself contingent. It is a contingent fact that ‘the external world *does not* exist’ is neither true nor useful information and can in no way impact upon how I choose to act. If the world were instead like the fictional world of *The Matrix*, such information would be highly useful; it would then be the kind of proposition that could be known. Significantly, if ‘The external world does not exist’ was known, ‘If there is something that I know, then I know that the external world exists’ would be false. Since, as mentioned, it is only a contingent fact that our world is not like the world of *The Matrix*, it is therefore only a contingent fact that this proposition is true.

We have here two *a priori* knowable contingent truths: ‘If there is something that I know, then I know that the external world exists’ and ‘The external world exists’. The former is more controversial, but an explanation of its apriority is available. Indeed, for each uncontroversial example of the contingent *a priori* a more controversial example can be constructed; more controversial, but boasting the aforementioned explanation. Specifically, from an uncontroversial φ we can construct ‘If there is something that I know, then I know that φ ’. Moreover, it feels as though we ought to be able to move from the controversial to the uncontroversial; we should be able to deduce φ from the fact that we know φ . If this can be done, we will have produced indirectly our explanation of the a priority of φ . Unfortunately, the move is not simple, since on

²⁴Ibid.

²⁵Craig himself resists this suggestion.

Craig's account knowledge is no longer factive; it is possible for someone to be an objectively good informant and be mistaken. It is feasible, however, that closer reflection on the relationship between φ and 'If there is something that I know, then I know that φ ' will prove fruitful. In any case, it is a relationship worth noting.

References

- [1] Bealer, G. (1999) 'The A Priori', Greco & Sosa (eds), *The Blackwell Guide to Epistemology*. Oxford: Blackwell Publishing.
- [2] Bostock, D. (1988) 'Necessary Truth and A Priori Truth', *Mind*, 97/387: 343–379.
- [3] Chalmers, D.J. (1996) *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.
- [4] Craig, E. (1990) *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. Oxford: Oxford University Press.
- [5] Descartes (2008) M. Moriarty (ed), *Meditations on First Philosophy*. Oxford: Oxford University Press.
- [6] Donnellan, K. (1977) 'The Contingent A Priori and Rigid Designators', in *Midwest Studies in Philosophy 2*: 12–27.
- [7] Evans, G. (1979) 'Reference and Contingency', in *Monist* 62: 160–89.
- [8] Fitch, G.W. (1977) 'Are There Contingent A Priori Truths?', in *Journal of Critical Analysis* 6/4: 118–123.
- [9] Gettier, E. (1963) 'Is Justified True Belief Knowledge?', in *Analysis*, 23/6: 121–123.
- [10] Kaplan, D. (1979) 'On the Logic of Demonstratives', in *Journal of Philosophical Logic* 8: 81–89.
- [11] Kripke, S. (1980) *Naming and Necessity*. Cambridge: Harvard University Press.
- [12] Plantinga, A. (1974) *The Nature of Necessity*. Oxford: Oxford University Press.
- [13] Plato (1996) J.H. McDowell (ed), *Theætetus*. Oxford: Clarendon Press.

- [14] Putnam, H. (1981) *Reason, Truth and History*. Cambridge: Cambridge University Press.

Has Davidson rejected conceptual relativism?

Anne Deng

London School of Economics

In this paper I will argue that Davidson fails to reject the existence of an incommensurable conceptual scheme. This is because although this is an ontological issue, only half of his argument is ontological, while the other half is epistemological. By an ontological issue or argument, I mean that it is concerned with the existence of a certain notion – in this case, it is whether incommensurable conceptual schemes *exist*. By an epistemological issue or argument, I mean that it is concerned with our knowledge on that notion – in this case, it is whether we can *know of* or *identify* such a scheme. I will establish this ontological–epistemological contrast by stating both halves of Davidson’s argument, explain this contrast with the distinct notions employed (definition and identification), and consider the objection that such a pattern is oversimplified.

Before diving into Davidson’s argument, I will first explain the notion of conceptual schemes and illustrate his strategy in refuting the existence of an incommensurable conceptual scheme. For Davidson, a conceptual scheme is a system of *concepts* that enables us to organise experience and form beliefs.¹ Two incommensurable conceptual schemes are thus schemes that contain concepts which cannot be defined in terms of the concepts of each other. Davidson’s first step in rejecting the existence of an incommensurable conceptual scheme is to accept the doctrine that associates a conceptual scheme with a set of languages that are inter-translatable. Based on this doctrine, he provides the following formulation of incommensurable conceptual schemes and aims to reject it – “two people have different conceptual schemes if they speak language that fail of intertranslatability”.²

There are two strategies available to reject this claim. Firstly, there is the usual strategy in which one shows that the failure of inter-translatability is in fact not a sufficient condition for the existence of incommensurable conceptual schemes; i.e. even if two sets of languages are not inter-translatable, they are still not associated with incommensurable conceptual schemes. Secondly, there

¹Davidson [1].

²Ibid. p. 7.

is the strategy in which one shows that the antecedent is simply false; i.e. there do not exist two languages that fail inter-translatability. Although this technically does not disprove the conditional in question, it counts as an adequate rejection, because it leaves the truth of the conditional vacuous.

What Davidson does is exactly the two strategies suggested above. He points out that there are two cases of failure of translatability – one complete in which no significant range of sentences can be translated to sentences in another language, and one partial in which some range can be translated and some cannot. In the case of complete failure of translatability, Davidson adopts the second strategy in which he shows that completely untranslatable languages do not exist. In the case of partial failure of translatability, he adopts the first strategy in which he operates on the basis that two sets of languages are partially untranslatable and aims to show that they are not associated with incommensurable conceptual schemes. In the following I will show that the first argument is ontological whereas the second one is epistemological.

II

In the case of complete failure of translatability, Davidson rejects its existence by showing that translatability is built into the definition of language. He provides the following formulation of languagehood – “something is a language . . . if it stands in a certain relation (predicting, organizing, facing or fitting) to experience (nature, reality, sensory promptings)”.³ He argues that there are two relations that might work: languages either *organises* its empirical content, i.e. they give empirical data a formal structure, or they *fit* it, i.e. they predict or account for an empirical fact. In terms of what is being *organised* or *fit*, he also offers two possibilities: language either stand in a relation with reality itself or our experience about it.

Let us first consider the possibility that languages *organise* reality or experience. Davidson points out that we can only organise something if it consists of other elements.⁴ For example, if we are to organise a closet, we would not be dealing directly with the closet itself but the shoes and shirts inside. This means that for languages to organise reality or experience *in any way*, they must at least contain concepts that individuate the *same* objects or familiar principles that individuate the *same* experience. However, if this were true, the languages would cease to be completely untranslatable. Davidson hence con-

³Ibid. p. 13

⁴Ibid.

cludes that the relation of languages *organising* their empirical content cannot be used to define languagehood without making assumptions about at least partial translatability.

The same point can be made about the relation that languages fit reality or experience. Davidson argues that this relation is essentially equivalent to the truth of the sentences in that language, and argues that the definition of truth requires translatability using Tarski's theory of truth as our best intuition.⁵ For Davidson, the fact that we recognise sentences like “‘snow is white’ is true if and only if snow is white” as trivially true can only entail the extension of the concept of truth for English. Hence, if one wishes to establish a concept of truth for all languages, one's first step is always “for every sentence *s* of [any language] \mathcal{L} , a theorem of the form ‘*s* is true if and only if *p*’ where ‘*s*’ is replaced by a description of *s* and ‘*p*’ by *s* itself if \mathcal{L} is English, and by a translation of *s* into English if \mathcal{L} is not English”.⁶ That is, the definition of truth for all other languages must make use of the notion of translation into English (or any other language in which the theorist is working). Hence the relation of language fitting reality or experience cannot be established without translatability either, and overall, languagehood cannot be defined without the notion of translatability. That is, languages that completely fail inter-translatability do not exist.

III

Having explained Davidson's argument on complete failure of translatability, I will now turn to his argument on partial failure of translatability. As I aim to show, this is where his argument becomes epistemological in the sense that it no longer deals with the existence of but our knowledge about incommensurable conceptual schemes – more on it later. As partial failure of translatability (i.e. partial success of translatability) already assures us that the language under consideration is indeed a language, all there is left to reject is that it is associated with an incommensurable conceptual scheme, i.e. the associated scheme consists of concepts that are undefinable by the concepts of our own scheme, or any other given scheme.

Davidson first points out that in order to identify such concepts, we must interpret the sentences in that language without making assumptions about

⁵Ibid.

⁶Ibid. p. 17.

“shared meanings, concepts or beliefs”.⁷ Here, he employs his own theory of radical interpretation based on *the attitude of accepting as true*, and argues that it is impossible to interpret a language in such a way. This is because holding a sentence true is the vector of two interdependent forces – “a workable theory of meaning” and “an acceptable theory of belief” and both meaning and belief requires a certain assumption of the concepts one uses.⁸ For example, if you see a ketch sailing by but your companion says ‘look at that handsome yawl’, you are faced with two ways of interpreting his sentence: firstly, by assuming what he *means* by an ‘yawl’ and a ‘ketch’ is the same as you, you may attribute the false *belief* to him that a yawl has sailed by; however, if he has good vision, you may decide that ‘yawl’ and ‘ketch’ have different *meanings* to him, and attribute him a correct *belief* – which is the same as yours, just in different terms. This interdependency means that it is impossible to even start to interpret a language without assuming either meanings or beliefs. Therefore, we are unable to *identify* any undefinable concepts in a supposedly incommensurable conceptual scheme, because to do so is to interpret the sentences in its associated language, and interpretation requires us to make certain assumptions about meanings or beliefs, which is essentially to make assumptions about concepts themselves.

Unlike the argument on complete failure of translatability, this argument is clearly not ontological but epistemological. This is because instead of the existence of an incommensurable conceptual scheme, i.e. the existence of undefinable concepts, this argument deals with the *identification* of such concepts. To make this distinction clearer, consider another pair of philosophical issues – whether reality exists (to which *idealists* answer no) and whether we can know of that reality (to which *scepticists* answer no). The argument on complete failure of translatability is of the first kind, because it shows that untranslatable languages are not possible and hence that it does not exist. The argument on partial failure of translatability, however, is of the second kind, because it shows that we cannot *know of* an incommensurable conceptual scheme. This idea is in fact present in Davidson’s own paper where he comments on his second argument – “we have found no intelligible basis on which it can be said that schemes are different”.⁹ That is, there is so far no coherent theory that enables us to *identify* two schemes as incommensurable.

⁷Ibid.

⁸Ibid. p. 18.

⁹Ibid.

IV

Having stated the ontological–epistemological contrast between Davidson’s argument on complete and partial failure of translatability, I will now explain this contrast with the different notions employed. In order to do this, I will first make a distinction between the definition of X and that of the identification of X , where X can be anything that ranges from abstract concepts such as justice to actual objects such as a table. This distinction is significant because although people talk of them interchangeably in everyday context, neither of them entails the other. For example, although philosophers have not yet *defined* the notion of causation, we are still able to *identify* a causal relation when we see one.

I propose that the argument on complete failure of translatability is ontological, because it is concerned with the failure of a coherent *definition*. Here, just as one deduces from the logical contradiction of a square circle that it doesn’t exist, by showing that an untranslatable language is undefinable due to the inherent necessity of translatability in the definition of languagehood, Davidson makes the ontological claim that languages that completely fail intertranslatability do not exist.

The argument on partial failure of translatability, however, differs from this argument in at least two ways. Firstly, it assumes that we already have a *definition* for conceptual incommensurability, because the purpose of Davidson invoking the notion of radical interpretation is to apply that definition and identify undefinable concepts. If he didn’t already have a definition in mind, he would be analysing the notion of conceptual incommensurability in the same manner that he treats the notion of languagehood in his first argument. Secondly, it further concerns with the notion of *identification*, because Davidson in this argument shows that we cannot know of undefinable concepts and hence incommensurable conceptual schemes. Based on these two considerations, the distinction between *definition* and *identification* hence nicely explains the fact that this second argument is epistemological rather than ontological.

Before I conclude, I will consider a possible objection against my argument. One might argue that the pattern in which the ontological–epistemological contrast corresponds to the definition-identification distinction is oversimplified, and there often is not a perfect match. For example, the fact that X can be identified might act as the evidence that X exists, despite the fact that X has not yet been coherently defined. This is the case with a number of significant philosophical notions, e.g. causality, laws of physics, etc. However, this objection misunderstands my argument in the sense that I do not require a perfect match between the two distinctions. All I am arguing for is that the *failure*

of establishing an account of identification makes the *negative* epistemological claim that we cannot know of incommensurable conceptual scheme, and it is only the *failure* of establishing an account of definition that makes the *negative* ontological claim that such a scheme does not exist. Whether those two distinctions match in the case of success in establishing an account of definition or identification and making a positive ontological or epistemological claim does not concern the current discussion.

Overall, I have argued that Davidson fails to reject the existence of incommensurable conceptual schemes, because the non-existence of an incommensurable conceptual scheme is an ontological issue, whereas only one of his arguments is ontological (the one on complete failure of translatability) while the other is epistemological (the one on partial failure). His first argument makes the ontological claim that completely untranslatable languages do not exist and hence an incommensurable conceptual scheme cannot be associated with a completely untranslatable language. His second argument makes the epistemological claim that we are unable to *identify* an incommensurable conceptual scheme, and it has no implication on whether such a scheme really exists.

References

- [1] Davidson, D. (1974) 'On the Very Idea of a Conceptual Scheme', in *Proceedings and Addresses of the American Philosophical Association* 47: 5–20.

MA Philosophy Durham University



MA Philosophy

Course structure

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

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

Modules offered are subject to change.

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

Career Opportunities

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

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

Entry requirements

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

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

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

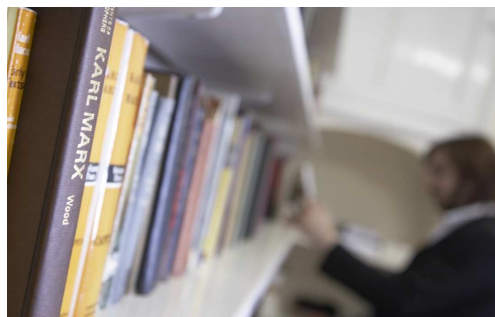
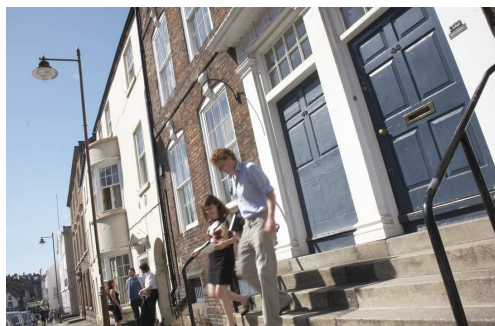
How to apply

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

Enquiries

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

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



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

Issue 11 (1)
Winter Conference 2017

*Explanatory Gap and Psychological
Explanation*

Benedetta Romano

*Conceiving 'Ideally' and its implications
for the Modal Argument*

Ahmad Jabbar

Against Atomism

Gareth Pearce

*Post-hoc Memory Erasure as a Response
to the Kane-Widerker Objection*

Harry Sherwood

Kant's Epistemology of Testimony

Anni Raty

Solipsism in the Tractatus

Sam Davies Udina

*A Host of Uncontroversial A Priori
Knowable Contingent Truths*

Oliver Hurcum

*Has Davidson rejected conceptual
relativism?*

Anne Deng

