15 Schriften der Fakultät Geistes- und Kulturwissenschaften
der Otto-Friedrich-Universität Bamberg

Metaphysics or Modernity?

Contributions to the Bamberg Summer School 2012

Edited by Simon Baumgartner, Thimo Heisenberg and
Sebastian Krebs

University
of Bamberg
Press

15 Schriften aus der Fakultät Geistes- und Kulturwissenschaften der Otto-Friedrich-Universität Bamberg

Schriften aus der Fakultät Geistes- und Kulturwissenschaften der Otto-Friedrich-Universität Bamberg

Band 15

University
of Bamberg
Press
2013

Metaphysics or Modernity?

Contributions to the Bamberg Summer School 2012

Edited by Simon Baumgartner, Thimo Heisenberg and Sebastian Krebs

University of Bamberg Press
2013

Bibliographische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Informationen sind im Internet über http://dnb.ddb.de/ abrufbar

Herstellung und Druck: DigitalPrint Group, Nürnberg
Umschlaggestaltung: University of Bamberg Press, Andra Brandhofer

http://www.uni-bamberg.de/ubp/
Titelfoto: Rupert Ganzer (http://www.flickr.com/photos/loop_oh/)

ISSN:1866-7627
ISBN: 978-3-86309-180-4 (Druckausgabe)
eISBN: 978-3-86309-181-1 (Online-Ausgabe)
URN: urn:nbn:de:bvb:473-opus4-46761

A note from the editors

This collection of essays originated in the Summer School 'Metaphysics or Modernity?', which was held at the University of Bamberg in August 2012. Designed as a forum for graduate students in philosophy, the Summer School brought together a highly diverse group of young academics who – more often than not – came from utterly different schools and traditions of thought.

This pluralism is reflected in the pages of this book. While the volume is roughly divided into two halves – one with a more *historical* focus, the other with a more *systematic* focus – the reader will find an unusually wide array of topics and questions treated here. Since the aforementioned pluralism was one of the main strengths of our Summer School, this is something in which we take much pride.

Neither this collection of essays nor the Summer School itself would have been possible without the generous support of both the Philosophy Department at the University of Bamberg and the German Academic Exchange Service (DAAD). In particular, Christian Schäfer and Christian Illies have spent numerous hours discussing our project with us, navigating administrative difficulties and giving valuable advice or encouragement. To them – as well as to all our colleagues who were willing to publish their work in this collection – we owe heartfelt thanks.

Simon Baumgartner, Sebastian Krebs and Thimo Heisenberg

Contents

Modernity, Metaphysics, and the Re-Invention of Philosophy – A Short Historical Introduction

Prof. Christian Schäfer
(University of Bamberg)

There seemed to be something like a broad consensus in 20th century philosophy that metaphysics corresponded to a 'premodern', and hence, obsolete, way of thinking. This consensus reflected a more general anti-metaphysical sentiment in the sciences, social life, and theology. But even if we leave aside the fact that this purported consensus presupposes a dubious way of looking at epoch-making and history altogether and a somewhat poor philosophy of history, and although in contemporary philosophy this consensual view seriously appears to crumble: the consensus on metaphysics' incompatibility with modern thought still shows what can be considered a continuing standard notion concerning metaphysics and the self-conception of modern man. As a matter of fact, metaphysical thinking is widely believed to have been supplanted by a more effective or less gullible way of thinking sometime during the 17th or 18th century. The result of this historical development in philosophy is that every philosopher who, in spite of the prevailing consensus, wants to do metaphysics today has to meet the challenge posed by the anti-metaphysical sentiment in philosophy.

The challenge consists of two components. First, one's metaphysical position must address the anti-metaphysical consensus (an awkward task, since it is difficult to define it neatly), and account for what enables one to pursue metaphysics in the face of it. In other words, one must explain why one thinks that the longstanding criticism of metaphysics does not invalidate metaphysical thinking, but should have had a cathartic effect on doing metaphysics. Second, and partly dependent on the reply to the first problem, contemporary metaphysics has to tackle an even more demanding task: it has to explain and vindicate its concept of philosophy, for the question of metaphysics has always been one concerning the way philosophy as a whole is conceived of.

Both these challenges converge at a certain point. Every criticism of metaphysics presupposes a certain concept of the metaphysical and its value for philosophy, and in turn this is liable to influence and, as the case may be, to re-define what is conceived of as 'metaphysical' (i.e., the material and the formal object of metaphysics). In this introduction, I therefore evaluate how these two problems, or rather clusters of problems, present themselves in the philosophical currents that in the course of centuries shaped the consensus on metaphysics as 'premodern' thinking. I thus consider the question of 'metaphysics and modernity' from a historical viewpoint in order to give an idea of the decisive currents of the anti-metaphysical consensus. The first problem concerns understanding the prevailing patterns of anti-metaphysical sentiment during the last three hundred years or so and the difficulty of discerning which conception of metaphysics underlies these patterns. The second problem concerns the related issue of the transformations of metaphysical conceptions and conceptions of metaphysics resulting from the anti-metaphysical sentiment in the history of Western philosophy. Due to space constraints, I limit myself to what I believe are the three most influential philosophical 'schools' of anti-metaphysical criticism, and after explaining their basic ideas, I attend to the question of how each of them contributed to a specifically 'modern' criticism of metaphysics.

1. Transcendentalising the subject and 'de-transcendentising' the metaphysical

Philosophers have traditionally distinguished metaphysics into 'general' and 'special' metaphysics. The latter speculates about the nature of God ('being as such' in a definitive acceptation), the soul, providence etc., whereas the former considers the basic structure of reality ('being as such' in a fundamental acceptation); questions about substance, essence, categories, the status of universals and individuals, etc., are paramount here. From the 16th century on, Protestant thinkers in particular (admittedly relying on longstanding, but, until then, rather underrepresented theological traditions) tended to view 'special metaphysics' as falling below acceptable standards of 'rationality.' They

held that religious truths were attainable by scripture alone, as the 'formal principle' of Protestantism maintained, and by faith alone, as the 'material principle' of Protestantism stated. This was part of their effort to 'purge' religion from philosophy and from natural theology. Accordingly, in a move that can be considered the first serious attempt to confine metaphysics to the 'premodern', philosophers like Christian Wolff accented the difference between 'special' and 'general' metaphysics. For him, 'metaphysics' meant *metaphysica specialis* and was devoid of any entitlement to philosophical rationality, whereas 'general metaphysics' he renamed 'ontology.' This re-conceptualisation kept general metaphysics – alias 'ontology' – philosophically afloat for some time yet. As a consequence, the term 'metaphysics' was assigned almost exclusively to private, mostly irrational convictions, mystical insights, and the sensitive issues touching on scriptural truths and other supposedly rationally unaccountable matters. This left 'metaphysical thinking' to such seemingly outré domains as alluded to by Woody Allen's famous joke "I was thrown out of N.Y.U. my freshman year for cheating on my metaphysics final, you know. I looked within the soul of the boy sitting next to me."

The dichotomy between general and special metaphysics and the disparagement of the latter culminated in what is generally called the 'subjective turn' of modernity. From the 16th century onward, a novel inwardness characterized modern piety and considerably influenced the estimation of metaphysics. Ontology, on the other hand, awaited its own subjective turn that would reduce the transcendent to the transcendental, and transfer the basic structuring of reality to psychological preconditions. Although he had not intended to abolish metaphysical thinking altogether, it was Kant who brought about this subjective turn with his 'critical philosophy.' In his *Prolegomena to Any Future Metaphysics That Will Be Able to Come Forward as Science*, he reduces the question "how is metaphysics possible at all (as a science)?" to the question "how is cognition from pure reason possible?". Kant's 'preliminaries' or 'prerequisites' for all future metaphysics contend that the 'thing in itself' (i.e., reality considered as such) cannot be an object of philosophical scrutiny. According to Kant, we must begin to conceive of reality *quoad nos*, as we 'phenomenologically' experience it and as we are capable of critically understanding it. In Kant's transcendental

dismissal of the thing in itself, *anything* that 'transcends' the subject belongs to a rationally unaccountable 'beyond', and thus, the objects of general metaphysics – as the objects of special metaphysics before – were 'sent to kingdom come,' to the realm of the rationally unaccountable. Accordingly, after Kant, the problems of 'general metaphysics' or of what once was ontology shifted almost exclusively to questions about (human) cognition, the mind-body problem, human autonomy and the like. In order to be philosophically acceptable, the traditional field of 'general metaphysics' dissolved into an analysis of the subject's cognitive makeup – its irreducibility, its range, and its possible function as a source of normativity.

2. Logic, empiricism, and logical empiricism

Whatever the definition of 'metaphysics' may be, its standard acceptation certainly seems to presuppose that metaphysics describes and deals with certain basic structures subjacent to reality which a) are rational and logically coherent but b) transcend or precede the realm of human experience and the phenomenological. This is in perfect agreement with the traditional main topics of metaphysical thought, e.g. universals, categories, God, questions of essence, potentiality, aprioristic knowledge, or human liberty. In its heyday during the first half of the 20th century, the criticism of metaphysics zeroed in on both these aspects of traditional metaphysics. Logical empiricists, some of them emerging from neo-Kantian currents, others from post-Hegelian schools, held and explicated a positivist view that constrained all knowledge to experientially verifiable empirical perception (as its 'material principle': by empiricism) on the one side, and everything logically analyzable (as its 'formal principle': by logic alone) on the other. Metaphysical statements, logical empiricists assert, do not pertain to either of these two realms or sources of knowledge (in the wider ambit of these thinkers philosophers like Wittgenstein or Russell, along with the mainstream of analytic philosophy, held occasionally different, but quite comparable views). Accordingly, logical empiricism saw the role of philosophy predominantly, if not solely, in explaining how to reconcile logical (or, as the case may be, mathematical) truths and methods with

the empirical sciences and their results. It was in contention to this confinement of philosophy and to the basic tenets and discussions of logical empiricism that the new interest of analytical philosophy in metaphysics arose. Saul Kripke's writings on modal logics are but one famous example of how logical empiricism's critique of metaphysical thinking has been challenged successfully in the 20th century.

3. Continental 'existentialism'

The project of re-inventing philosophy to the detriment of metaphysics is even more palpable in existential thought, stemming from Nietzsche and – via Kierkegaard and Heidegger – culminating in the likes of Camus and Sartre. If in the context of anti-metaphysical thought, in post-Kantian 19th century mainstream philosophy and logical empiricism 're-invention' meant to 're-conceive' philosophy in an attempt to re-define it in such a way as to show that it had nothing or little in common with the dark ages of metaphysical philosophy, Nietzsche launched, and Heidegger intensified, an attack on metaphysical thinking where 're-inventing' meant to 're-discover' philosophy, assuming that philosophy had been led astray since the times of Plato, or even before, and alleging that the different forms and systems of metaphysical thinking were exactly what accounted for the wrong track of philosophy in history. According to them, traditional metaphysics had been a 'new' kind of philosophy that had betrayed 'true philosophy.' They called for a rediscovery of what philosophy had originally been – and should be again. With its famous adage – "existence precedes essence" – the existentialist movement – despite the serious doctrinal differences between its exponents – advocated a reorientation of philosophy that abjured traditional metaphysical concepts such as essence and the category of the universal. It instead tried to ground our understanding of Being from a non-systematic consideration of human existence and its imponderabilia. In existentialism's anti-metaphysical creed, it is the individual's consciousness (to avoid the term 'intellect' or 'reason') that produces values and determines meaning in an absurd world devoid of it and any rational coherence. According to Heidegger, for example, a

'fundamental ontology' is called for that substitutes traditional 'transcendentals' with what he calls 'existentals.' He advocates that philosophy illuminate our pre-reflective understanding of the world and of the human condition by meditating on metaphysically 'unspoiled' everyday meanings of concepts such as time, being, death, possibility, and thinghood. He claims that his philosophy offers a more primordial approach to the question of being (*Sein*) through an analysis of being human (*Dasein*) than classical metaphysics, and other existentialists did likewise, everyone in his own fashion.

This brief historical survey, streamlined as it may be, demonstrates at least one thing: The criticism of metaphysics always presupposes a 'thick' conception of the metaphysical, and the forms of 'new metaphysics' have a symbiotic relationship with anti-metaphysical criticism evident in the recent history of Western philosophy. A historical survey like the one given above can clarify the quality of this criticism up to a certain point: Criticism of metaphysics has a long tradition in philosophy. But generally, it has always been either more radical or less general than the kind of criticism that has culminated in the standard dismissal of metaphysics as 'premodern.' Consider the Pyrrhonist attack on metaphysical thought that resulted in a skeptical attitude towards rational explanations, or the nominalist 'anti-metaphysics within metaphysics' in the dispute over universals. Hardly any of this is found in the criticisms of metaphysical thinking in modernity. The problem of 'metaphysics and modernity' is historically considered a meta-philosophical problem, one that is sui generis and aims at banning and abolishing, rather than quarreling or arguing with, metaphysical thinking. Forms of criticism that declare that metaphysics is a premodern way of thinking claim that metaphysics is a way of reasoning that does not apply to our (that is, currently prevailing) way of thinking, and that it does not comply with our (that is, prevailing 'normal scientific') acceptance of the rational. They dismiss metaphysics as non-philosophical. Therefore, in contrast to the attack on metaphysical thinking launched by, say, Ockham or Hume, the different contemporary approaches that we label 'criticism of metaphysics' do not aim at particular metaphysical tenets. This is true regardless of the fact that every type of criticism of metaphysics focuses on a certain type of metaphysics, as it seems to target a specific form or variety of

metaphysical theory, or stems from a specific doctrinal controversy with a particular school of metaphysical thinking: For example, the Protestant pre-Kantian thought took on scholastic Aristotelianism and the existentialists on traditional metaphysical essentialism – let alone the fact that in their alleged 'demolition of metaphysics' Kant's critical philosophy and other idealists' philosophical projects resemble Neoplatonic thinking and its *very* metaphysical, basic convictions to a great extent. But this should not distract us from the fact that all attempts to dismiss metaphysics as a 'premodern' way of thinking, despite their particular quarrelling with certain forms or schools of metaphysical thought, nonetheless rely on the conviction that philosophy as a whole has to be redefined and that within that definition of philosophy there is no place left for metaphysical thinking, neither in its acceptance as special metaphysics nor as general metaphysics.

John Locke's Contemporaries' Reaction against the Theory of Substratum

Mihretu P. Guta
(Durham University)

The goal of this paper is to critically examine the objections of John Locke's contemporaries against the theory of substance or substratum. Locke argues in *Essay* that substratum is the bearer of the properties of a particular substance. Locke also claims that we have no knowledge of substratum. But Locke's claim about our ignorance as to what substratum is, is contentious. That is, if we don't know what substratum is, then what is the point of proposing it as a bearer of properties? This question underlies the criticism Locke's contemporaries raise against the notion of substratum. In section I, I lay out the context for Locke's theory of substratum by pointing out his main motivation in proposing his theory. In section II, I give a brief analysis of the theory of substratum. In section III, I discuss the objections of Locke's contemporaries against the theory of substratum.[1] I focus on Edward Stillingfleet, Lee Henry, G. W. Leibniz and John Sergeant. In section IV, I conclude that there is no warrant to dismiss Locke's theory of substance.

1. The Need for Substratum

Locke's proposal of the theory of substratum aims at answering one key question. That is, do properties (e.g., color) or qualities need bearers? As we shall see, Locke answers this question affirmatively. But how does Locke go about the task of accounting for the bearer of properties? How does he describe it? Does the supposed property bearer have its own nature, i.e., its own identity via which we come to know what it is? Does Locke have a uniform and uncontentious way to characterize his theory of substratum? Such questions are still hotly debated and disagreements over what constitutes the right answer to them are far from over.[2]

[1] For dialectical purposes, in section I, I consider the first objection against Locke's theory of substratum. In doing so, I will pave the way for my discussion of the other objections in section III.

[2] Both David Hume and George Berkeley dismissed Locke's theory of substratum (see: Lowe 1995: 83-87). But contemporary Locke scholars proposed a number of interesting ways to amend the contentious aspect of Locke's substratum. Due to space limitations, we

What is a substance? Locke answers this question by pointing out the role of senses and reflections (*Essay* Book II, chap. XXIII).[3] Locke claims that the mind is furnished with a number of simple ideas or qualities, which are found in 'exterior things.' But what is the source of these simple ideas? Locke claims that such simple ideas are obtained via the senses. Furthermore, for Locke, the mind on its own operations is capable of noticing the unity of qualities, i.e., such qualities 'go constantly together' in experience (*Essay* II, XXIII. 1). So such unity observed in simple ideas, is often attributed as belonging to one thing, even describing it under one name. But for Locke, what is being understood as one simple idea is rather a combination of many ideas (Ibid.). However, here we need to keep in mind that even if Locke considers an idea as subjective mental phenomenon (e.g., see *Essay* II, VIII. 8), his use of the term 'idea' is not always fixed. That is, Locke also uses the term 'idea' to refer to a *quality* of a subject existing external to the mind which produces a particular idea in our mind (see further Lowe 1995:19-22).

Locke tells us that the notion of the unity we observe in simple ideas forces us to ask a question. That is, what enables these simple ideas, i.e., qualities of a physical object to stay in unity? More precisely, what underlies such unity? Initially, Locke's answer for this question may come across both as arbitrary and ad hoc. For example, Locke claims that simply because we cannot make sense of how qualities or simple ideas can subsist by themselves, we tend to assume that something grounds them or supports them. As Locke states:

> ...not imagining how these ideas can subsist by themselves, we accustom our selves, to suppose some *substratum*, wherein they do subsist, and from which they do result, which therefore we call *substance* (*Essay* II, XXIII. 1).

---

will not be able to discuss the proposed solutions. However, towards the end of this essay, I will say something very brief in regards to the implications of the contemporary philosophers' attempt to amend Locke's substratum. While there are so many issues one can raise in relation to Locke's theory of substratum, for present purposes, the focus of this essay is narrow and is limited only to some of the issues Locke's four contemporary critics raise.

[3] Unless indicated otherwise, all references of the *Essay* are from: John Locke, *An Essay Concerning Human Understanding*, ed. Peter H. Nidditch (Clarendon: Oxford University Press, 1975).

Here some take Locke's remarks at face value and think that Locke's proposal of *substratum* as the bearer of sensible qualities is just a place holder. That is, its significance lies only in helping one make sense of the underlying ground for the unity of qualities whether or not the idea of substance itself is real. Commenting on the above quote, Alexander Campbell Fraser suggests:

> The expressions 'not imagining how'—'we accustom ourselves to suppose' seem to refer to our idea of substance to 'imagination' and 'custom,' instead of finding it implied in the very intelligibility of experience; for although 'custom' may explain our reference of such and such 'simple ideas' or qualities to such and such particular substances, it does not show the need in reason for substantiating them, in order to conceive that they are concrete realities (in Fraser, v. 1: 390, footnote 3).

If Locke's skepticism (or seeming agnosticism) towards the reality of substance is substantiated, then so much is worse for Locke. Lowe remarks that, among other things, the reason why Locke's account of the idea of substance generated more controversies than any other topic we find in Locke's *Essay* has to do with its implications for theological thought. This was most importantly in relation to the accounts of God's nature and the immortality of the soul.

So, doctrines such as (1) transubstantiation, i.e., a view that bread and wine transforms into the body and blood of Christ; and (2) the Trinity, i.e., the Father, Son and Holy Spirit while distinct persons, yet share the same indivisible divine nature, are rooted in the idea of substance (Lowe 2005:59). So the only plausible way that seems available to maintain the intelligibility of the idea of substance, as Lowe remarks, "would be to declare it innate" (Ibid.). In light of this, Lowe further remarks that for the religious establishments of Locke's time, Locke's empirical based understanding of the idea of substance was seen as a slippery slope down the road of atheism (Ibid.). This is because, *inter alia*, since God is taken to be a substance, to endorse Locke's account of substratum would require us to confess ignorance about our knowledge of God himself. But such confession (for the religious establishment of Locke's day) of our ignorance of the knowledge of God is nothing short of moving in the direction to embrace atheism.

But is the above fear of the dangerousness of Locke's view of substance justifiable? Details aside, one way to answer this question is to

look at Locke's response to one of his main critics, the Bishop of Worcester Edward Stillingfleet. As we recall, Locke employs some suspicious sounding phrases when he talks about substratum or substance: "...we accustom ourselves to suppose..." For Stillingfleet, such phrases came across as unacceptable on the basis of their implications for theological thought as briefly discussed earlier. So, in his third *Letter* to Stillingfleet, Locke clarifies his use of the phrase 'supposing' claiming that it should not be taken as a ground to label him as being skeptical of the reality of substance. In his *Letter* to Stillingfleet, Locke makes it clear that since we cannot conceive the existence of qualities *per se* without being substantiated, it follows that there must be something we call *substance* that underlies them (see Fraser, footnotes *Essay* II, XXIII. 1:390-391).

In this case, Locke is claiming to be a realist about substance ontology. Taken this way, the suspicion we put forth earlier as to whether or not Locke is a realist about substance ontology seems to lack any ground. Thus, we can say that the notion of 'substance' for Locke is not just a place holder after all. But as we shall see, such a positive characterization of Locke as a realist about substance ontology does not seem to weaken the objections his critics raise against his theory of substratum.

2. What is Substratum?

Locke claims that we have no idea of the notion of *pure substance in general,* or substratum. All that we can say with respect to substratum is something that supports the qualities that produce simple ideas in us, yet we can give no further analysis for it. For Locke, *pure substance* is simply something 'one knows not what supports'. As Locke states:

> If any one should be asked, what is the subject wherein Colour or Weight inheres, he would have nothing to say, but the solid extended parts: And if he were demanded, what is it, that Solidity and Extension adhere in, he would not be in a much better case, than the Indian...who, saying that the world was supported by a great Elephant, was asked, what the Elephant rested on; to which his answer was, a great Tortoise: But being again pressed to know what gave support to the broad-back'd Tortoise, replied, something, he knew not what (*Essay* II, XXIII. 2; also cf. Bk. II. Ch. Xiii § 19).

Here Locke's point is that inquiry into what exactly substratum is, is not an open ended one. Even if, we know well that qualities (e.g., colour, weight) have owner or subject that instantiates them, the qualities themselves have no role to play by way of revealing the 'identity' of the thing that underlies them. So 'substratum', despite its key role in underlying qualities, is epistemically inaccessible to us. Hence, we cannot keep on asking endlessly what substratum is. In light of this, Locke suggests that the best way to end our curiosity, to get to the bottom of the identity of substratum, is by confessing ignorance.[4] Perhaps, here Locke's emphasis on ignorance could be taken as a deterrent to unnecessary explanatory regress. But despite such *prima facie* benefit, Locke's own insistence on the unknowability of *pure substance* turns out to be less illuminating. We will return to this discussion in section III. But insofar as Locke is concerned, he sums up his theory of substratum as follows:

> The idea then we have, to which we give the general name Substance, being nothing, but the supposed, but unknown support of those Qualities, we find existing, which we imagine cannot subsist, *sine re substante*, without something to support them, we call that Support *Substantia*; which, according to the true import of the Word, is in plain *English*, *standing under*, or *upholding* (*Essay* II, XXIII. 2).

[4] That said, however, it is important to keep in mind that for Locke, *the general idea of substance* is not made of via a process of combining many simple ideas, which eventually lead to the formation of complex ideas. By contrast, the abstract or general idea of substance is formed only by the process of abstraction and hence it refers to a mental process. As Locke puts it, "The use of Words then being to stand as outward Marks of our internal *Ideas*, and those *Ideas* being taken from particular things, if every particular *Idea* that we take in, should have a distinct Name, Names must be endless. To prevent this, the Mind makes the particular Ideas, received from particular Objects, to become general; which is done by considering them as they are in the Mind such Appearances, separate from all other Existences, and the circumstances of real Existence, as Time, Place, or any other concomitant *Ideas*. This is called *ABSTRACTION*, whereby *Ideas* taken from particular Beings, become general Representatives of all of the same kind; and their Names general Names, applicable to whatever exists conformable to such abstract *Ideas*....Thus the same Colour being observed to day [sic] in Chalk or Snow, which the Mind yesterday received from Milk, it considers that Appearance alone, makes it a representative of all of that kind; and having given it the name *Whiteness*, it by that sound signifies the same quality wheresoever to be imagin'd or met with; and thus Universals, whether *Ideas* or Terms, are made" (*Essay*, II, XI. 9). See Berkeley's objection to Locke's doctrine of abstraction in the introduction to the *Principles*.

3. Reactions against Locke's Theory of Substance

Other contemporaries of Locke also followed Stillingfleet's footstep in critiquing Locke's view of substratum. Here is how Henry Lee begins by explaining his own version of substance before he critiques Locke's:

> This Name of Substance we give to any thing whose Existence we conceive independent upon every thing else, and in which several Properties or Qualities are united and combined. And this, as old as it is, is taken to be a perfect Definition of Substance in General: because hereby the Mode of its Existence is distinguishable from that of Qualities or Properties: namely, its Existence does not depend upon any other created Substance or any Quality or Property. Whereas Qualities or Properties do depend upon one Substance or other, and have no qualities united in them. (II.23.1: 110).

Here Lee's analysis of the notion of substance is Aristotelian in its tone. Lee thinks that the relation that exists between substance and properties or qualities is asymmetrical, that is not equal. To make sense of the asymmetry Lee draws here, we can do no better than return to Aristotle himself. Aristotle *inter alia*, holds that substances (*ousia*) are ultimate subjects of predication and ontologically independent. In the *Categories*, Aristotle argues that primary substances have ontological priority, i.e. while other things depend on them for their existence, the converse is not true. Aristotle claims that primary substances such as individual men and horses are subjects that ground the existence of other nonsubstantial things such as qualities and quantities. For Aristotle, the primary substances are ultimate subjects of which other things are predicated but which are not themselves predicated of anything else (1028^b 36-37). In the *Categories,* Aristotle does not treat primary substance as complex bodies although he does treat primary substances as complex bodies in his *Metaphysics* (see further: Mary Louise Gill 1989: chap.1).

Because of such a shift in Aristotle's thinking, there is a debate on whether complex entities, i.e. entities with combination of matter and substantial form are primary substances. Yet Aristotle's own commitment to the primacy of substance over stuff still stands. This line of thought has been rigorously defended in recent years (see e.g., Lowe 1999: chap. 1; Lowe 1998: chaps. 6-9; Michael Loux 1998: chap. 3). Since for Aristotle, stuff has only potentialities as opposed to actualities, stuff

fails to be basic. To say that rock is stuff of a building is to say that rock has only a potential to become a building. By contrast, actuality is prior to potentiality (Book Θ 8, 1049b 18-25). Since substance is actuality, it follows that substance is prior to stuff. But Locke rejects the Aristotelian notion of *substantial forms*. But for now, the details do not concern us.[5] Here the point is that when Lee refers to his construal of substance as 'old', his approach is broadly Aristotelian. If this is correct, then what is Lee's point here?

Before we answer this crucial question, it is important to bear in mind that Aristotle's discussion of substance ontology is inextricably linked with his metaphysics. In fact, it is hardly an overstatement to say that we can only arrive at a proper understanding of Aristotle's substance ontology, if we take Aristotle's metaphysics seriously which underlies Aristotle's overall approach in his investigation of the nature of reality. In *Metaphysics* Γ.1, Aristotle tells us that the object of the investigation of *first philosophy* is not as limited as some special sciences, which only tend to focus on their respective area of interest. For example, mathematics focuses on things that are countable and measurable. But for Aristotle, metaphysics is a universal science that studies being *qua* being.[6] But what does Aristotle mean by 'being qua being'? As S. Marc Cohen remarks, Aristotle's description of 'the study of being qua being' does not imply as if there is a single subject matter—being qua being—which is under investigation. Instead the phrase, 'being qua being' involves three things: (1) a study, (2) a subject matter (being), and (3) a manner in which the subject matter is studied (qua being), (Cohen 2012: sec. 1). As Cohen further points out, Aristotle's study does not focus on some recondite subject matter identified as 'being qua being'. Rather it is a study of being. In other words, for Aristotle, first philosophy studies beings, in so far as they are beings (Ibid.).

So as Michael Loux argues, metaphysics considers things as existents and tries to specify the properties they exhibit in so far as they are beings. The chief goal of metaphysics is not just to grasp the concept of being but also to grasp general concepts such as identity, difference,

[5] For an excellent discussion on 'substantial form', see Oderberg 2007: 65-71.
[6] Here the term 'qua' comes from Latin and it means 'in so far as' or 'under the aspect'.

similarity, and dissimilarity that apply to everything that there is. In light of this, Loux further remarks that central to Aristotle's metaphysics is the description of what Aristotle refers to as categories. These are the most general kinds under which things fall. So the business of a metaphysician is to identity those general or highest kinds. But the task does not end here. That is, a metaphysician is also supposed to specify the features unique to each category. The upshot of engaging in such an activity provides us with a map of the structure of all there is (Loux 1998: 3-4).[7]

So for Aristotle, metaphysics is the study of the fundamental structure of reality.[8] Insofar as Aristotle is concerned, a metaphysician's success in the investigation of the nature of reality depends on whether or not his/her approach is a realist one, which is to say that whether or not one accepts the existence of a mind-independent reality. For Aristotle, a mind-independent reality is the starting point of his metaphysical theorizing. If this is true, then I fully agree with Kit Fine when he says, "I take Aristotle's primary concern in his metaphysical and physical writings to be with the nature of reality rather than the nature of language. I am rarely tempted, when Aristotle appears to be talking about things, to construe him as saying something about words..." (Fine 1996: 83-84).[9]

---

[7] See also, other contemporary neo-Aristotelian advocates of the traditional metaphysics: Lowe 1998; ibid. 2002; ibid. 2006; Oderberg 2007; Hoffman and Rosenkrantz 1997; Schaffer 2009: chap. 12; and Tahko (ed.) 2012.

[8] Similarly, in his recent *Writing the Book of the World*, Ted Sider describes metaphysics as an inquiry into the fundamental structure of reality (Sider 2012). However, it is somewhat troubling to note that there is no mention of Aristotle in Sider's book. One wonders what justifies Sider's decision to do so.

[9] But the hitherto brief description of the traditional metaphysics was rejected by David Hume, Immanuel Kant and in modern times, by post Hume-Kant thinkers such as W.V. Quine. Most notably, a serious attack was launched against the traditional metaphysics by the 1930s and 40s Vienna Circle logical positivists, such as Rudolf Carnap and others. As is well known, logical positivists promoted a movement that restricts the source of genuine knowledge to what can only be empirically verifiable (see e.g., Uebel 2012). Though, at present, the 'verification principle' largely has fallen out of favor, it is safe to say that its spirit is still around. Recently Huw Price in his essay entitled, "Metaphysics after Carnap: The Ghost Who Walks?", argued that metaphysics [traditional] is 'as dead, or at least deflated, as Carnap left it' (Price 2009: 322). Here Price's remark is not only very strong but it is also entirely unfounded in light of the work of notable contemporary neo-Aristotelian advocates of traditional metaphysics (see e.g., the representative sample I gave above under footnote # 7). One can confidently say that Aristotelian metaphysics is back on the stage,

In light of the foregone considerations, we are better prepared now to see Lee's point when he refers to his construal of substance as 'old'. In this regard, Edwin McCann in his essay 'Locke's Theory of Substance under Attack!', points out that Lee is basically distinguishing what he calls the 'old' notion of substance from Locke's 'new' one. In doing so, Lee wants to show that Locke's notion of pure substance obscures the old notion of substance. Lee's main goal then is to diagnose Locke's mistake (McCann 2001: 95). As Lee puts it:

> He [Locke] would have a clear notion of pure Substance (that is) abstracted from all Properties or Qualities whatever. I answer, there is no such Substance in the whole World. Every Substance has some Qualities or other: A Spirit its Thinking, Space its Expansion, Body its Solidity. And then how can any man have a distinct Notion of that, from which you suppose all its Properties (by which it should be distinguish'd) separated or abstracted?...there is really no such Substance in the World abstracted from all Qualities, nor no Qualities abstracted from all Substances (II.23.3:110-111).

Here the biggest worry Lee expresses against Locke's notion of pure substance seems to be that of the divorce Locke introduced between pure substance and property. As we recall, for Locke, substratum, though it is a property bearer, is not itself knowable. So Lee's point is that once we divest pure substance of all its properties (via which we come to know what pure substratum is), the result is to find ourselves in the dark. What then is the implication of such ignorance of what pure substance is? If we let Lee himself answer this question, he might say that if a substance is devoid of a property that defines its identity, then the very existence of such a substance can be called into serious question. Lee

---

contra to Price's claim and all those who are sympathetic to his remarks. But I don't intend to argue for this claim here for that would take us too far afield. That said, however, unlike Hume who dismissed both Aristotelian metaphysics in general, and his substance ontology in particular (see *Treatise* Book I part IV), Kant was a friend of substance ontology, which he went on defending it. As he insightfully remarks, "...the necessity with which this concept of substance forces itself upon us, we have no option save to admit that it has its seat in our faculty of a priori knowledge", (Kant, *Critique of Pure Reason*, B.6). So whatever reservations one might have about Locke's characterization of substance, it should not be a reason to reject Aristotelian substance ontology. One should keep in mind that despite his departure from the Aristotelian tradition, the roots of Locke's own conception of substance ontology can be traced back to Aristotle. Despite some knee jerk reactions of some modern philosophers, Aristotelian substances continue to occupy a central place in our discussion of identity, agency and causation, *inter alia* (see e.g., Ruggaldier 2006: chap.3; also see Lowe 1998: chaps. 4-7).

also might say that we have no epistemic ground to talk about a substance that is abstracted or separated from its property. This is simply because for Lee, we cannot conceive a substance apart from its property on the one hand and the instantiation of properties without their bearer on the other. In light of this, Lee remarks:

> But if he [Locke] expects we should answer what a Substance is, 'tis not kindly done to suppose, first, the Question is unanswerable, as it is, if we strip it of all of its Properties and Qualities, and then ask us, what it is? (II.23.4:111).

Along similar lines with that of Lee, Leibniz also in his *New Essays* expresses serious doubts against Locke's characterization of pure substance. Here are Leibniz's own remarks:

> THEO. If you distinguish two things in a substance - the attributes or predicates, and their common subject - it is no wonder that you cannot conceive anything special in this subject. That is inevitable, because you have already set aside all the attributes through which details could be conceived. Thus, to require of this 'pure subject in general' anything beyond what is needed for the conception of 'the same thing' - e.g. it is the same thing which understands and wills, which imagines and reasons - is to demand the impossible; and it also contravenes the assumption which was made in performing the abstraction and separating the subject from all its qualities or accidents. The same alleged difficulty could be brought against the notion of being, and against all that is plainest and most primary. For we may ask a philosopher what he conceives when he conceives 'pure being in general'; since the question excludes all detail, he will have as little to say as if he had been asked what 'pure substance in general' is. So I do not believe that it is fair to mock philosophers, as your author does [xiii.19] when he compares them to an Indian philosopher who was asked what supported the world, to which he replied that it was a great elephant; and then when he was asked what supported the elephant he said that it was a great tortoise; and finally when he was pressed to say what the tortoise rested on, he was reduced to saying that it was 'something, he knew not what'. Yet this conception of substance, for all its apparent thinness, is less empty and sterile than it is thought to be. Several consequences arise from it... (II.23.2:218).[10]

[10] Gottfried Wilhelm Leibniz also criticized Locke's remarks on the doctrine of innate ideas (*New Essays on Human Understanding*, 1740). Here 'innate idea' refers to the notion that the basic components of human knowledge and understanding are not rooted in sense experience. Rather such items of knowledge and understanding (e.g., necessary truth) were imprinted in the human mind. So human beings, so to speak are born with such innate or primary ideas to which universally all humans assent to. For Leibniz, as it is true for Rene Descartes, such construal of innate ideas is perfectly justifiable (see, e.g., John Harris in 'Locke on Human Understanding,' I.C. Tipton (ed.), 1977: chap. II; Mackie 1976: chap. 7). But Locke thinks that such blind acceptance of innate ideas stifles honest inquiry into how we acquire our knowledge thereby encouraging dogmatism as well as obscurantism. So

It is clear that Leibniz also thinks that divesting the concept of substance of its attributes is problematic. This is because for Leibniz, we can only come to grasp the very nature of substance via its qualities or properties. In light of both Lee's and Leibniz's insistence on the importance of ascribing a property to a substance (obviously the move Locke resists), the question remains: what sort of property can we ascribe to a substance? Here essential property and accidental property come to mind. Roughly speaking, an object is said to have an essential property if it cannot exist without it. For example, being extended is an essential property of a house, without which a house cannot continue to exist. By contrast, an object is said to have an accidental property if an object can lose it without ceasing to exist. For example, suppose that John is sad now. It is possible for John to have lacked the property of being sad now. So being sad is an accidental property of John. For present purposes, I will put aside complicated and detailed discussions that surround these two domains of properties, i.e. essential and accidental (see further Lowe 2002, chap. 6; Code 1986; Fine 1992; McGinn 2000; Plantinga 1974).

In light of this, it might be the case that both Lee's as well as Leibniz's complaints against Locke's characterization of substratum or pure substance as 'something we know not what' seem to have targeted on the loss of an essential property of a substance. However, here we may argue that Locke does not necessarily deprive a substance of its nature. Rather Locke only insists that we have no knowledge of the nature of substance. But such response does not satisfy Locke's critics. This is because, since by Locke's own account we have no idea of what the nature of substance is, it remains unclear what entitles us to say anything substantial about it.

So the heart of Lee's and Leibniz's charge against Locke's characterization of substance seems to be that Locke deprived pure substance of all its content. But some disagree with such a conclusion.

Locke dedicates Book I of his *Essay* to refute the doctrine of innate ideas before he returns in Book II of his *Essay* to the discussion of substratum and particular substances. For more information on Locke's take on innate ideas and the controversies surrounding it, see e.g., Lowe 1995: chap. 2 and 2005: chap. 2; J. Gibson 1917: chap. 2, § 6; J.W. Yolton 1956; R.I. Aaron 1955; D.J. O'Connor 1967: 39-40, Greenville Wall in Locke on Human Understanding, I.C. Tipton, ed., 1977: chap. I, Margaret Atherton in Locke, Vere Chappell, (ed.), 1998: chap. II and J.L. Mackie 1976: chap. 7.

For example, McCann claims that both Lee and Leibniz mischaracterized Locke's characterization of substance. McCann thinks that Locke did not leave us with an empty conception of substance nor did he subtract all properties and qualities from it, as both Lee and Leibniz claim. Furthermore, for McCann, Locke does not rob the notion of substance of all content. This is because, as McCann argues, Locke's characterization of the idea of substance as self-subsisting property bearer of objects (serving as their common subject) exactly matches Lee's own characterization of substance as the 'old' conception of substance (McCann 2001: 96). However, McCann's defense of Locke's conception of substance appears to be overly simplistic.

If indeed Locke's conception of substance exactly matches that of Lee's, then that means that Locke takes an Aristotelian conception of substance without reservation. But that is indeed not the case. Locke clearly does not endorse some core notion of substance in the Aristotelian tradition in general and scholastics, in particular. For example, although Locke calls such things as man, horse, sun, etc., as particular substances, he does not consider them to be substances in the fundamental ontological sense. McCann's misunderstanding here arises from his unjustifiable inference that because Locke grants that substance is a property bearer, it follows that Locke's conception of substance is not devoid of content. The real issue here is not the lack of consensus or awareness of the fact that Locke clearly takes pure substance to be a property bearer. Rather the real issue has to do with what pure substance itself is, on its own ground apart from being a property bearer. Does it have its own nature or identity, if it does what is it? How do we come to know it? So, only knowing *pure substance* as a property bearer *per se* is not a substitute as McCann seems to think for settling such fundamental questions.

So, contrary to McCann, it is not wrong to say that Locke left us with an empty conception of substance. For example, Jonathan Bennett argues that many philosophers rightly point out that Locke's notion of *pure substance* or *substratum* is impossible or intolerable. But why is that? Bennett thinks that the answer concerns with conceptual emptiness. By this Bennett meant that since it is thought that because *substratum* has to be the bearer of all the qualities, it must be therefore, in itself, bare or unqualified in some problematic way (Bennett 1998: 131

in Vere Chappell, ed.). But as Bennett further points out the defender of substratum theory is inevitably accused of requiring content in something that had been stripped of all content (Ibid.: 132).

McCann's move in trying to bring Locke's conception of *pure substance* closer to the traditional understanding of substance turns out to be inconsistent with McCann's own suggestion about what he thinks Locke is up to, in advancing his theory of substance. McCann points out that Locke's main agenda in Book II of the 23rd chapter of his *Essay* has been understood in terms of (1) Locke giving us a positive theory of substance more in line with the traditional view of substance; and (2) Locke giving us a negative theory of substance, which is taken to imply that he is entirely rejecting his theory of substance (e.g., see *Essay* II.13.17-20 & II.23.2). But McCann suggests his own alternative: (3) the 'no-theory' theory of substance. As McCann puts it:

> Locke is giving an account of an idea we have, the idea of substance in general, which is derived from simple ideas given in sensation and reflection but which, given the vagaries of this derivation, is hopelessly and irremediably confused and obscure. An idea this confused and obscure cannot be used to explain anything, including the six phenomena....Locke's realization of, and indeed insistence upon, this fact explains why it is that he never, either in the *Essay* or in the correspondence with Stillingfleet, claims to explain anything in terms of that idea. It is an idea we've arrived at *faute de mieux*, serving mainly to mark our inability to conceive how qualities exist by themselves by supposing they exist in *something*, something, given the circumstances, we know not what (McCann 2001: 94-95).

Here by 'six phenomena' McCann is referring to six features attributed to the theory of substance which he himself recognizes can be traced back to Aristotle. In a nutshell, these features are (a) predication relation that is, certain terms are predicated of other in an asymmetrical fashion. We can say, the 'flower is red' but it is wrong to say 'The red is a flower'; (b) substances can exist on their own but the same cannot be said of modes or properties. That is, only the later need the former for their instantiation; (c) substances unify various sensory experiences (e.g., colors, shapes, feels, smells) to hang-together; (d) substances play the role of identifying or individuating individual objects. For example, a substance can persist via qualitative change while maintaining its numerical identity over time; (e) substances have a special sort of unity in that a substance is *ens per se*, i.e., a being through itself, which is

radically different from mere aggregates, which are *entia per accidens*, i.e., an entity that exists accidentally; (f) the theory of substance also needs to tackle the question of how many basic kinds or fundamental kinds of substance there are and how each basic kind differs from the other (Ibid.: 88-90).

What is the problem with McCann's (3) above?[11] The problem is not with McCann's overall characterization of Locke's theory of substance. Certainly one can ask whether or not Locke's theory of substratum is explanatorily adequate. But the issue here is whether or not Locke is successful in the way he characterized his theory of *substratum*. Leaving that aside for now, McCann's (3) is inconsistent with his earlier remarks that Locke's theory of substance 'exactly' matches both Lee's as well Leibniz's characterization of substance. How so? Given McCann's (3) Locke's theory cannot explain the six features of substance McCann himself lists as shown above in (a)-(f). That means that Locke's theory of substance hardly matches the account given by Lee as well as Leibniz.

This is because the Lee-Leibnizan account is perfectly consistent with the six features pointed out by McCann, whereas Locke's theory of substance fails to match up with the six features. This is precisely because Locke's substratum or pure substance is so obscure and confused whose features we cannot tell (if such a task is even possible to undertake in the first place). If this is correct, then my early remarks regarding the Lee-Leibnizan account of substance being Aristotelian in its spirit stands their ground. If I am right about this, it follows that McCann's claim that Locke's conception of substance is equally rich in its content as that of Lee-Leibnizan turns out to be by its own standard both inconsistent and groundless.

Moreover (3) can be given an ambiguous reading. Basically (3) does not tell us anything different that cannot be achieved either by adopting one of the other strategies, i.e., (1) or (2) as stated above. But the ambiguous reading I have in mind here can be spotted between (2) and (3). As it stands, one can understand (3) as Locke rejecting his theory of substance, in which case the meaning of (2) is in place. On the other

[11] As we recall, we already made (with some qualification) a positive case for (1). But we rejected (2) on the basis of claiming that Locke is a realist about substance ontology. In the above quote, McCann mistakenly thinks that the idea of substance is derived from simple ideas given in sensation and reflection. See footnote # 4 for Locke's own view.

hand, one can also understand (3) as implying that Locke is not rejecting his theory of substance, instead he is simply stating the difficulties surrounding it to spell out what it is, in which case we settle for (3). But it is hard to tell which meaning fixes McCann's 'no-theory' theory of substance. In light of such considerations, therefore, option (1) seems to be the best alternative to adopt.

Finally, the other contemporary of Locke worthy of mentioning is John Sergeant. Sergeant begins his discussion of Locke's 23rd chapter of Book II as follows:

> While I perused Mr. *Locke*'s 23th Chapter, of the *Idea of Substance,* I was heartily grieved to see the greatest Wits, for want of True Logick, and thro' their not lighting on the right way of Philosophizing, lay Grounds for Scepticism, to the utter Subversion of all *Science;* and this, not *designedly,* but with a *good Intention,* and out of their Sincerity and Care not to *affirm* more than they *know.* He fancies that the Knowledge of *Substance* and *Extension* are absolutely Unattainable. Now, if we be altogether Ignorant what *Substance* or *Thing* means, we must bid Adieu to all Philosophy, which is the *Knowledge of Things,* and confess that we talk all the while of *we know not what:* And, if we be invincibly Ignorant of what *Extension* is, farewell to all the Mathematical Sciences; which, (those that treat of *Number,* or Arithmetick, excepted) do all of them presuppose our Knowledge of *Extension,* and are wholly *grounded* on that Knowledge (Reflection, 1697, 13.22-23: 238).

Sergeant certainly gives credit to Locke for exercising an epistemological modesty by not pretending to know more than he does. Yet for Sergeant such epistemological modesty comes with a price. That is, suspending judgment in matters such as substance and extension creates a fertile ground for skepticism. Hence, the ultimate consequence of such skepticism shakes up the very foundation of our knowledge of things. So, Sergeant offers his own version of substance:

> Now, as we can consider the Thing *according* to its *Modes* or *Accidents,* so we can have *another* Notion or Consideration of the *Thing* as to *its own self*· abstracting from all these former Considerations; or a Notion of the *Thing,* (not according to any Mode it has, but) precisely according to its *Thingship* (as we may say) or Reality; that is, in order to *Being;* or (which is the same) we can consider it precisely and formally as an *Ens, Res, Substance* or *Thing;* and all we can say of it, *thus consider'd,* is, that *'tis capable to be actually.* For, since we see Created Things have *Actual Being,* yet so that they can *cease to be;* all that we can say of them, (thus consider'd) is, that they are *Capable to be.* Besides, since we see they *have* Being, were this *Actual Being* or Existence *Essential* to them, they would *be of themselves,* and so *could not but be;* and, consequently, must *always be;* which our common Reason and Experience contradicts; in regard we

> know them to have been *made;* and we see many of them daily *Produced,* and others *Corrupted* (Reflection, 1697, 13.22-23: 239-40).

Compared to the Lee-Leibnizans' account of substance, Sergeant's account is different. Sergeant spells out his notion of substance in terms of 'capacity to exist'. In this regard, McCann points out that Locke responded to Sergeant's definition of substance in two marginal comments on Sergeant's own *Solid Philosophy*. Locke's first comment reads: "All which amounts to no more but that Substance is something, which is what Mr. L. [Locke himself] says" (p. 241 as quoted in McCann 2001:97). The other comment Locke gave concerns direct criticism of Sergeant's own conception of substance:

> If the Idea of Substance be capacity to exist, then Accidents are Substances for they are capable to exist. If it be as J.S. puts it here and else where a thing capable to exist, then his Idea of substance or thing will be this, that a thing is a thing capable to exist, which as much clears the point as if he should say an [accident?] is an accident capable to exist, or a man capable to exist (p. 244 as indicated in McCann 2001: 98)

Here I agree with McCann's remark that Locke's criticism of Sergeant's conception of substance is effective. Locke is certainly right to say that if the criteria for something to be a substance is based on the capability of whether that thing exists or not, then anything that is capable to exist can be given a status of substance, a notion that goes against what we have discussed so far. But where I still disagree with McCann is when he says, "Locke is criticizing Sergeant for departing from the traditional notion of substance'' (Ibid.: 98). Here my disagreement with McCann is based on the earlier claim that Locke is not committed to the traditional notion of substance to the extent McCann claims. Thus, at the least, it remains unclear whether or not Locke's point in criticizing Sergeant's construal of substance is motivated by Locke's commitment to the traditional understanding of the theory of substance. Although the four contemporaries of Locke have their own independent line of thoughts, what unifies them all against Locke is their insistence on the inadequacy of Locke's characterization of *substratum* or *pure substance* in general as 'something we know not what.'

The question remains: where does all this leave Locke's theory of substance? Is Locke's theory of substance misguided? I personally think

not. As Sergeant rightly noticed, Locke's insistence on the unknowability of substratum is motivated by his epistemological modesty (i.e., by not pretending to know more than he does). Yet as we recall from sections I & II, Locke's main goal in proposing *substratum* does seem to be primarily metaphysical in nature. That is, Locke on a purely ontological ground seems to have realized that properties or qualities necessarily need some sort of bearer for them to exist. Hence, there cannot be free floating qualities (cf. Lowe 1995: 76-77). More importantly, as Lowe remarks, when Locke characterizes substratum as 'something we know not what', he (Locke) seems to be implying that it may have a nature which may be known to other beings such as angels and God (e.g. *Essay* II, XXIII. 6; Letter to the Bishop of Worcester, p. 28 as quoted in Lowe, 2000:507). If so, the objections Locke's immediate critics raised against his theory of substance are not insurmountable.

In light of such and similar other considerations, contemporary philosophers such as C.B. Martin, E.J. Lowe, M. R. Ayers, Jonathan Bennett, J.L. Mackie, Margaret Atherton, and Martha Brandt Bolton proposed various solutions to provide a defensible framework for Locke's theory of substance. Though the amendments these philosophers suggest differ from each other, they all agree that Locke's theory of substance is not a result of careless conjecture and thus must not be dismissed. For reasons already indicated, looking at these solutions is not our present concern (see for details, Martin, 1980; Lowe, 2000 & 2005; Ayers, in Tipton, I.C., 1977; Mackie, 1976; Bennett, Ayers, Atherton and Bolton, in Vere Chappell 1998). That said, however, one thing we can certainly say is that with some modification, Locke's theory of substance can be made free from the problems that beset it.

4. Conclusion

In this short paper, we have looked at what Locke's motivation was in proposing the theory of substratum. We have also looked at some of the problems that beset the theory of substratum as pointed out by Locke's immediate critics. Yet Locke's insistence on the unknowability of substratum stems from his epistemological modesty. If so, the best that

can be said about the merit of the objections raised by Locke's contemporaries is that at the least, the objections show that Locke's theory of substratum is contentious. But such objections in no way prevent us from modifying Locke's account of substratum (as contemporary philosophers attempt to do). Therefore, I conclude that, as things stand, there is no warrant to dismiss Locke's theory of substance.[12]

Literature

Ackrill, J.L. (Trans.). *Aristotle's Categories and De Interpretatione*. Oxford: Clarendon Press, 1963.

Aaron, R.I. *John Locke*. Oxford: Oxford University Press, 2nd edition, 1995.

Berkeley, George. *Principles of Human Knowledge and Three Dialogues between Hylas and Philonous*. London: Penguin Books, 1988.

C.B. Martin. "Substance Substantiated." *Australian Journal of Philosophy* 58 (1980): pp.3-10.

Chappell, Vere (Ed.). *Locke*. Oxford: Oxford University Press, 1998.

Chalmers, David, Manley David and Wasserman Ryan (Eds.). *Metametaphysics: New Essays on the Foundations of Ontology*. Clarendon: Oxford University Press, 2009.

Code, Alan. "Aristotle: Essence and Accident." in *Philosophical Grounds of Rationality,* edited by R.E. Grandy and R. Warner. Oxford: Clarendon Press, 1986.

Cohen, S. Marc, "Aristotle's Metaphysics." *The Stanford Encyclopedia of Philosophy* (Summer 2012 Edition), edited by Edward N. Zalta, URL = <http://plato.stanford.edu/archives/sum2012/entries/aristotle-metaphysics/>.

Corradini, Antonella, Galvan, Sergio and Lowe, E. J. (Eds.). *Analytic Philosophy without Naturalism*. London: Rutledge, 2006.

Fine, Kit. "Essence and Modality." June 1992; revised February 1993, URL =

http://as.nyu.edu/docs/IO/1160/essence.pdf accessed 1 May 2012.

——— "The Problem of Mixture." *Form, Matter, and Mixture in Aristotle,* edited by Lewis A. Frank and Bolton, Robert. Oxford: Blackwell, 1996.

Gibson, James. *Locke's Theory of Knowledge*. Cambridge: Cambridge University Press, 1917.

Gill, Mary Louise. *Aristotle on Substance: The Paradox of Unity*. Princeton: Princeton University Press, 1989.

Hoffman, Joshua and Rosenkrantz, S. Gary. *Substance: Its Nature and Existence*. London: Rutledge, 1997.

[12] I am very thankful for E.J.Lowe, Sophie Gibb, Marrku Keinänen and Thimo Heisenberg for constructive feedback on the earlier draft of this paper.

Hume, David. *A Treatise of Human Nature,* edited by L.A. Selby-Bigge. Clarendon: Oxford University Press, 1955.

Kant, Immanuel. *Critique of Pure Reason,* edited by Norman Kemp Smith London: The Macmillan Press, 1929.

Lee, Henry. *Anti-Scepticism: or, Notes upon each Chapter of Mr. Locke's Essay Concerning Humane Understanding. With an Explication of all the Particulars of which He Treats, and in the same Order. In Four Books*. London: Printed for Clavel and C. Harper, 1702.

Leibniz, G.W. *New Essays on Human Understanding*, edited and translated by Jonathan Bennett and Peter Remnant . Cambridge: Cambridge University Press, 1996.

Locke, John. *An Essay Concerning Human Understanding*, commentary by Alexander Campbell Fraser. Oxford: Clarendon Press, 1894.

——— *An Essay Concerning Human Understanding,* edited by Peter H.Nidditch. Clarendon: Oxford University Press, 1975.

Loux, Michael J. *Metaphysics: A Contemporary Introduction*. London: Rutledge, 1998.

Lowe, E.J. *A Survey of Metaphysics*. Oxford: Oxford University Press, 2002.

——— "Locke, Martin and Substance." *The Philosophical Quarterly* 50, no.201 (2000): 499-512.

——— *Locke*. London: London: Rutledge, 2005.

——— *The Four Category Ontology: A Metaphysical Foundation for Natural Science*. Oxford: Oxford University Press, 2006.

——— *Locke on Human Understanding*. London: Rutledge, 1995.

——— *The Possibility of Metaphysics: Substance, Identity, and Time*. Oxford: Clarendon Press, 1998.

Mackie, J.L. *Problems from Locke*. Oxford: Clarendon Press, 1976.

McCann, Edwin. "Locke's Theory of Substance under Attack!" *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*. 106, no. 1/2 (2001): 87-105.

McGinn, Colin. *Logical Properties: Identity, Existence, Predication, Necessity and Truth*. Oxford: Oxford University Press, 2000.

O'Connor, D.J. *John Locke*. Dover Books, 1967.

Oderberg, David S. (Ed.). *Form and Matter: Themes in Contemporary Metaphysics*. Oxford: Blackwell Publishers, 1999.

Oderberg, David S. *Real Essentialism*. London: Rutledge, 2007.

Plantinga, Alvin. *The Nature of Necessity*. Oxford: Oxford University Press, 1974.

Sergeant, John. *Solid Philosophy Asserted, against the Fancies of the Ideists, or, The Method to Science Farther Illustrated: with Reflexions on Mr. Locke's Essay Concerning Human Understanding*. London: Printed for Roger Clavil...Abel Roper...and Thomas Metcalf..., 1697.

Sider, Ted. *Writing the Book of the World*. Clarendon: Oxford University Press, 2012.

Tahko, E. Tuomas (Ed.). *Contemporary Aristotelian Metaphysics*. Cambridge: Cambridge University Press, 2012.

Tipton, I.C (Ed.). *Locke on Human Understanding: Selected Essays*. Oxford: Oxford University Press, 1977.

Uebel, Thomas, "Vienna Circle." *The Stanford Encyclopedia of Philosophy* (Summer 2012 Edition), edited by Edward N. Zalta, URL = http://plato.stanford.edu/archives/sum2012/entries/vienna-circle/.

Yolton, John W. *John Locke and the Way of Ideas*. Oxford: Oxford University Press, 1956.

——— *Locke: An Introduction*. Oxford: Basil Blackwell, 1985.

Metaphysics in the work of Charles Darwin

Bárbara Jiménez Pazos
(University of the Basque Country)

It is not hard to see how two visions of nature are intertwined in Darwin's Journal of Researches: one vision, the province of romantic authors depicting the sentiments awakened by certain landscapes, the other, the domain of natural scientists describing the world without reference to the aesthetic qualities of the scenery. Nevertheless, analyses of this double perspective in Darwin's work are relatively rare. Most scholars focus on Darwin, the scientist, and more or less ignore the aesthetic aspects of his work. Perceiving the gradual transformation of Darwin's world view, however, depends on analyzing the two different modes in which Darwin approached and perceived the world. While one can, on occasion, find commentaries on the beauty of the natural world in Darwin's early work, the passage of time produces a modification in the naturalist's manner of perceiving nature. This does not, however, mean that Darwin ceases to find beauty in nature; on the contrary, the disenchantment, in Max Weber's words, that Darwin's theory produces should not be understood in a pejorative, but rather in a literal sense. The theory of evolution, in effect, divests nature of its magical character and begins to explain it in terms of natural selection, according it, in the process a new and more intense attraction. In the present work, the metaphysical implications of this new vision of the world are analyzed through the eyes of its discoverer.

1. Darwin, religion and landscape

> Formerly I was led by feelings, such as those just referred to (although I do not think that the religious sentiment was ever strongly developed in me), to the firm conviction of the existence of God, and of the immortality of the soul [...] I well remember my conviction that there is more in man than the mere breath of his body. But now the grandest scenes would not cause any such convictions and feelings to rise in my mind. It may be truly said that I am like a man who has become colour-blind (Darwin 1887: 311-312).[1]

As demonstrated by, among others, Daniel Dennett's book *Darwin's Dangerous Idea* (Dennett 1995) or by the disputes in the United States

[1] All citations from *Life and letters of Charles Darwin. Including an Autobiographical Chapter* are from Volume I.

over Intelligent Design theory, analyzing the metaphysical implications of Darwin's theory remains relevant to philosophical debates. The quote that opens the present work refers to the relation between the emergence of religious sentiment, produced by the aesthetic perception of landscape, and the disappearance of the same caused by the growth in scientific knowledge about nature. One may, in turn, directly link Darwin's confession to a series of metaphysical questions surrounding Darwinism, such as religious beliefs (Creationism), the revision of received ideas with the aim of attributing meaning to human history and human existence, the aesthetic perception of nature, or the problem of the existence of God. Dennett, in effect, makes an oblique reference to these questions when he argues that Darwinism functions as a "universal acid" that corrodes all received metaphysical ideas.

There is little consensus regarding the relevance of theological questions in understanding the Darwinian theory of evolution. In describing aesthetic marvels, Darwin occasionally uses language that implies a foundation in religious belief. As Phillip R. Sloan points out, however, such language is "not religious in the traditional sense", as it is "devoid of references to 'God', 'creation', 'providence', 'design', or the other categories of traditional theology" (Sloan 2001: 261). On the whole, then, one can say that Darwin's writings contain metaphysical references, but display no clear connection to traditional theology.

After Darwin's voyage in the *Beagle* (1831-1836), the influence of Thomas Robert Malthus's writings gave him a new vocabulary that distanced itself from Humboldtian[2] nature. This period was of great

[2] Alexander von Humboldt's accurately taken nature data and his detailed landscape descriptions make the reader form an almost inevitable association between him and Charles Darwin, one of his followers. Both authors started research voyages to know in detail the natural world, and it is a remarkable fact that Darwin carried with himself a copy of Humboldt's *Personal Narrative* (1814) taken as a helpful reference for his own research. Darwin's *Journal of Researches* preserves Humboldt's descriptive style, which is characteristic for its adoption of two different ways of analyzing and describing the elements of nature. On the one hand, Humboldt adopted the Romantic Movement's delight for nature and, consequently, he also maintained the romantic descriptive style of nature's elements. This particular descriptive method is easily recognizable, since it shows the beholder's subjective point of view and, therefore, the feelings that some scenes or elements of nature arouse in the observer. On the other hand, one recognizes in Humboldt's texts the descriptive method that natural sciences use to analyze the

significance for Darwin. His understanding of nature became more complex and, prompted by the work of Malthus to conceive of nature as a selector, his naturalistic ideas began to acquire a direction slightly different from that of his predecessor Humboldt, though preserving traces of Humboldt's influence. As Sloan puts it:

> The synthesis of Darwin's complex body of readings and reflections finally took shape in the remarkable first drafts of his transformist theory in 1842 and 1844, the first texts to employ the concept of "natural" selection. In these texts we can also see the beginnings of the interplay of the "Humboldtian" and "Malthusian" conceptions of nature (Sloan 2001: 264).

Darwin's *Journal of Researches* (1839) clearly reveals this presence of two intertwined visions of nature: one is that of Romantic authors who want to convey the feelings that the observation of certain types of landscapes produces, and the other is that of natural scientists who describe the world without reference to the aesthetic qualities of the landscape. While one may characterize Darwin's early view of nature as pantheistic, one cannot say the same of the conception of nature present in the writings dating from the period of the *On Origin of Species* (1859), where the influence of Malthus is visible, and where the enumerative and schematic description of observations predominates. During the earliest years in which Darwin was engaged in formulating his theory, his writings, according to Sloan, display a conception of nature that resembles Spinoza's notion of *natura naturans* and *natura naturata*[3], and it is in those writings that one can perceive the link between the sense of the sublime in Darwin and the figure of God (with the understanding, as noted above, that this "God" is not the God of traditional theology). Connecting Darwin's ideas with the Spinozistic notion of nature, Sloan makes the important claim that in Darwin's writings,

> These [the complex forms and adaptations found in nature] are not contrivances imposed by an external creator on a passive material nature in the tradition of

---

environment. This descriptive style is formed by the objective information with which the observer aims to devise a complete view of nature, leaving aside the personal repercussion.

[3] One could refer to *natura naturans* as "active nature" (God) and to *natura naturata* as "passive nature" ("all that follows from the necessity of the nature of God"). These two states of nature can, on the one hand, be defined separately but, on the other hand, should be understood as two complementary elements (Spinoza 1989: 25-26).

> British natural theology. They are rather properties that emerge from the immanent constructive activities of nature itself (Sloan 2001: 268).

After all the references to theology, Sloan finds it necessary to qualify the term "religious":

> If we define "religious" motivations to involve some kind of belief in the transcendent, in an objective foundation of a moral order, and as the source of answers to the main questions of life and death, a constitutive and even religious significance in Darwin's appeal to "nature" is more plausible (Sloan 2001: 267).

One must analyze whether religious beliefs shaped the way in which the English naturalist formed an understanding of the natural environment. Sloan leans toward accepting that the transcendent had a certain influence on Darwin's ideas, causing Darwin to have a quasi-religious conception of nature:

> Darwin's "nature" was something more than a mere metaphysical premise [...] it was also a source of moral order for Darwin, not in the sense of a system displaying obvious design and contrivance, but as a lawful system on which one could rely for ethical norms, serving as the source and foundation for life. To this extent, we can affirm that cognitive premises of a quasi-religious nature do indeed play a significant role in Darwin's science (Sloan 2001: 269).

Seeing Darwin's conception of the quasi-religious nature as something that comprises a system of laws and ethical norms that establish the basic principle of life, one can conclude that this conception plays an important role in Darwin's scientific work.

David Kohn, in *Darwin's Ambiguity: The Secularization of Biological Meaning*, argues for a certain ambiguity in Darwin's metaphysics-related writings. On occasion, Kohn notes, references to the Creator coexist with a ridiculing of the special creation doctrine; similarly, the same letter may contain both the affirmation and the denial of God. Kohn sees, then, a naturalist with a dual character:

> So we continue to have these conflicting portraits: Darwin as conforming Victorian theist – the last of the natural theologians – and Darwin as religious radical who recovered the deistic tradition of the enlightenment and had a special role in establishing the independence of scientific naturalism and the secularization of the modern world view (Kohn 1989: 218).

According to Kohn, the important fact is that Darwin's search for a scientific theory of biological origins unfolds in a specific theological context (Kohn 1989: 223). The context appears to influence Darwin's thought and, according to Kohn, this influence can be seen during the first years in which Darwin developed his theory – although one can, at the same time, detect a certain materialism in his ideas. These two aspects of the naturalist add ambiguity to his texts.

> One can regard Darwin as a materialist, but this need not lead one to consider him an atheist, given that he saw no problem with affirming that the laws of nature may be established by God – an impersonal God, undefined and synchronized with nature itself. Darwin's work consists in explaining a natural order that implies the existence of a Creator (Kohn 1989: 238).

This double aspect is the reason that Darwin's thought is compared to Victorian Romanticism, which was also pulled in two different directions at once: the "lyrical materialism" (Beer 2009: 44) of William Wordsworth and the idealistic theology of Samuel Taylor Coleridge. Moreover, there is in Darwin a new approach to the sublime and the beautiful from the perspective of naturalism, as Kohn explains:

> Darwin can be shown to preach a naturalist reconciliation of the sublime and beautiful. When he lifts up the vision of a natural world created and finely balanced by selection, he captures the heightened religious emotions of a doxology and appeals to a spirituality dislocated by the Victorian crisis (Kohn 1989: 234).

The sublime, characterized by a powerful and destructive nature, is in diametric opposition to the beautiful, defined more by calm than by natural force. Wordsworth's search of a conciliation between the sublime and the beautiful is brought to completion in the work On the Origin of Species, where, according to Kohn, one encounters a Darwin who is no longer the young Humboldtian naturalist of the Journal of Researches, but rather a mature scientist capable of finding the balance between the beautiful and the sublime for which Wordsworth yearned (Kohn 1989: 235).

A close reading of the *Journal of Researchers* reveals the existence of two descriptive tendencies, two forms of approximating nature that appear to be intertwined, forming a stable equilibrium – and that are, for Darwin, complementary. Darwin achieves a new form of perceiving

nature, a form that is inevitably rooted in Romanticism, but that, little by little, distances itself from an *enchanted* (Weber) vision of nature, arriving at explanations of the way nature works that do not draw on the transcendental. Nevertheless, despite possessing an explanation of nature based on his theory of evolution, Darwin continues to perceive beauty in nature.

In my view, it is incorrect to claim that Darwin found himself on the fence at all times; rather, with the passage of time, Darwin refined his manner of observing nature. The increase of knowledge that resulted from his incessant observation of the environment caused Darwin to modulate his perception of the landscape. As James Paradis notes: "While Darwin's developing theories did not alter the appearance of landscape, they did ultimately alter what Darwin saw," (Paradis 1981: 105). Paradis locates the difference between Naturalism and Romanticism in the perception of nature, and comments that "as [Darwin's] concept of nature became increasingly intellectual and abstract, his representations became less traditional and emotional" (Paradis 1981: 107). I suspect that, contrary to Paradis's argument, knowledge intensifies the aesthetic experience, making it denser, more human, more serious, and more profound. This is exemplified in the last paragraph of Darwin's *On the Origin of Species*.

2. The entangled bank

Darwin closes the *Origin of Species* with the following words:

> It is interesting to contemplate a tangled bank, clothed with many plants of many kinds, with birds singing in the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to reflect that these elaborately constructed forms, so different from each other, and dependent upon each other in so complex a manner, have all been produced by laws acting around us. These laws, taken in the largest sense being Growth with Reproduction; Inheritance which is almost implied by reproduction; Variability from the indirect and direct action of the conditions of life, and from use and disuse: a Ratio of Increase so high as to lead to a Struggle for Life, and as a consequence to Natural Selection, entailing divergence of Character and the Extinction of less-improved forms. Thus, from the war of nature, from famine and death, the most exalted object which we are capable of conceiving, namely, the production of the higher animals, directly follows. There is grandeur in this view of life, with its several powers, having been originally breathed by the Creator into a few forms or into one; and that, whilst this planet has gone

> cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being evolved (Darwin 1876: 429).

According to Kohn (1997), the "entangled bank" metaphor does not pertain only to the last paragraph of *On the Origin of Species,* but is, rather, developed over the course of the *Journal*. Kohn argues that in the *Journal*, Darwin exhibits a variety of strong responses to natural scenery: in describing his experiences in Tierra del Fuego, he emphasizes decadence and desolation,[4] where, in describing the Brazilian forest, he sees grandiosity. Nevertheless, both scenes are conducive to the creation of the entanglement metaphor:

> The Darwinian fixation with entanglement, both in Tierra del Fuego and Brazil expresses a struggle towards the sublime that is rooted in Milton's language (Kohn 1997: 26).

It is worthy of note that Kohn, besides discussing the "entangled bank" metaphor, also traces the roots of the "wedging metaphor"[5] that appears in the *Transmutation Notebook* of 1838, describing it as the initial formulation of the thoughts expressed in the last paragraph of *On the Origin of Species*, where it is finally united with the "entangled bank". As Kohn puts it:

> My claim is that the birth of the entangled bank in 1859 was already foreshadowed in the *textual framing* of the 1838 wedging metaphor [...] the wedging metaphor and the entangled bank prove to be intimately related in the *Origin*. In the end, they are almost one [...] their commonality derives from Darwin's powerful attachment to his version of the romantic aesthetic (Kohn 1997: 40, italics in original).

Recalling the earlier discussion on the linkage of the sublime with natural force and of the beautiful with calm, Kohn adds one last thought regarding Darwin's metaphors. He argues that in both metaphors, one

---

[4] Kohn refers to the following quote: "Among the scenes which are deeply impressed on my mind, none exceed in sublimity the primeval forests undefaced by the hand of man; whether those of Brazil, where the powers of Life are predominant, or those of Tierra del Fuego, where Death and Decay prevail" (Darwin 1913: 533-534).

[5] The "wedging metaphor" derives from Darwin's comment about the necessity of finding a structure that could be adapted to change. As Kohn puts it: "The balance of death and destruction with life and growth – the oeconomic balance sheet of nature – finds its meaning in adaptative change" (Kohn 1997: 37).

can detect a kind of preparation of the textual frame, with Darwin inviting us to contemplate the natural landscape that he sees, so that he can then explain the workings of what he observes. In other words, one passes from the beautiful to the sublime within one thought.

The final fragment of *On the Origin of Species* displays aesthetic sensibility in a "disenchanted" sense, though one should not interpret that "disenchantment" as a negative process. One can see how Darwin continues to perceive beauty in nature ("endless forms most beautiful and most wonderful have been") despite having uncovered its workings.[6] In Darwin's case, there is no "disenchantment" in the Weberian sense – quite the opposite. Offering a theory that explains the unknown aspects of the evolution of species prompts Darwin to approach nature with new eyes, causing him to marvel at it even more than before he had discovered his theory of evolution through natural selection. George Levine (2008) terms this process "secular re-enchantment" and notes:

> The excitement that follows upon understanding the instincts that drive birds to migrate (and this requires no mystification or invocation of transcendental spirit), the astonishment that follows upon recognizing the overwhelming complexity of the eye's functioning (even despite the flaws in the mechanism that are clear evidence that there is no intelligent design behind the construction of the eye) [...] these and all the various knowledges that scientific study of nature and the human has been producing are elements of new forms of enchantment (Levine 2008: 28).

I find this approach to Darwin's perception of nature appealing, and therefore I venture to refine somewhat the claims of Paradis, discussed earlier: Darwin's theory of natural selection may not modify the appearance of landscape, but it does modify the representational instances (previous information, theoretical explanations, etc.) that always accompany human experience. This allows Darwin to see in the landscape that which the Romantics cannot see. In other words, Paradis's argument could be amended in the sense that while Darwinian evolutionism does not modify the "physical appearance" of the

[6] In fact, George Levine (Levine 2008: 31) argues that seeing Darwin's work as devoid of affection for the world is an incorrect way of interpreting it: "The tendency to understand Darwin's world as providing no affective or even rational compensation is [...] another of the 'misuses' – although perhaps an inevitable one – of Darwin".

landscape (a claim that is trivial, given that no theory can accomplish that), it does substantially reshape what the brain or mind sees. In other words, it does not alter what the eyes see, but it does change the human perception of the images on the retina; that is, it changes the interpretation of the landscape, the human sensations, the "apperception" (Leibniz) – and, finally, the internal experience, including the aesthetic experience.

3. Darwin's *Autobiography* and the loss of aesthetic taste

One of the aspects of the *Autobiography* that has received the most scholarly attention is Darwin's statement that he is losing his aesthetic taste. This loss begins with poetry, previously a source of pleasure to Darwin.

The naturalist notes in his autobiography that he used to read poetry. But he also confesses that he has completely lost the taste for it over the years: "Later in life I wholly lost, to my great regret, all pleasure of poetry of any kind" (Darwin 1887: 33). Darwin states that his taste for poetry lasted until he was thirty years old, but that since that age, poetry began to bore him to the point of nausea:

> Up to the age of thirty, or beyond it, poetry of many kinds, such as the works of Milton, Gray, Byron, Wordsworth, Coleridge, and Shelley, gave me great pleasure, and even as a schoolboy I took intense delight in Shakespeare, especially in the historical plays. I have also said that formerly pictures gave me considerable, and music very great delight. But now for many years I cannot endure to read a line of poetry: I have tried lately to read Shakespeare, and found it so intolerably dull that it nauseated me (Darwin 1887: 100-101).

Besides poetry, art and music had also given pleasure to Darwin. Yet in his old age, Darwin had so completely lost his taste for most art forms that he even felt that a part of his brain – the part related to the aesthetic (the higher tastes) – had atrophied; only some feeling for the beauty of scenery was left, and even that was reduced. Darwin finds this loss very strange:

> I retain some taste for fine scenery but it does not cause me the exquisite delight which it formerly did [...] This curious and lamentable loss of the higher aesthetic tastes is all the odder, as books on history, biographies, and travels (independently of any scientific facts which they may contain), and essays on all sorts of subjects interest me as much as ever they did. My mind seems to have

> become a kind of machine for grinding general laws out of large collections of facts, but why this should have caused the atrophy of that part of the brain alone, on which the higher tastes depend, I cannot conceive (Darwin 1887: 101).

Darwin's words are not simply a dispassionate acknowledgement of a loss of interest in poetry, painting, and music. Instead, Darwin conveys his grief at the loss, which he suspects may have further consequences:

> The loss of these tastes is a loss of happiness, and may possibly be injurious to the intellect, and more probably to the moral character, by enfeebling the emotional part of our nature (Darwin 1887: 102).

This "enfeebling of the emotional part of our nature" is what seems to worry Darwin most; this is hardly surprising given its importance for Darwin in previous years. In addition to the explicit emotion that he himself in the *Journal of Researches* described himself feeling, his son, Francis Darwin, recounts the childlike manner in which his father observed and admired flowers:

> I used to like to hear him admire the beauty of a flower; it was a kind of gratitude to the flower itself, and a personal love for its delicate form and colour. I seem to remember him gently touching a flower he delighted in; it was the same simple admiration that a child might have (Darwin 1887: 117).

This same admiration, that Francis Darwin considers childlike, can also be seen, though not in exactly the same form, in a passage in the *Journal of Researches,* where Darwin describes how his experiences in the Brazilian forest produced in him a feeling that there is more to human beings than the corporeal. By contrast, in the later *Autobiography* he describes himself as, in a way, color-blind. If the scenes of the Brazilian forest in years past had prompted in him a sense of belief in something more than the corporeal, with the passage of time, these scenes, so closely connected to a belief in God, ceased producing that effect in Darwin. The naturalist proclaims that the same scenes would no longer call forth in him the same sensations of the sublime, that is, of sentiments related to the divine.

It is possible that Darwin understood, or intuited, that the aesthetic sense, like the moral instinct, also has bio-evolutionary roots, so that the connection with the sublime, that aesthetic theories since Plato had affirmed, would be lost. It is known that Darwin lost at least the onto-

theological significance of the beauty of natural landscapes – that is, he no longer perceived them as *vestigia Dei* (traces of the Creator). This onto-theological perception of the landscape is inseparable from the traditional pre-Darwinist aesthetics: beauty as a manifestation of the sublime (divine) in man and in nature.

Various authors have contributed towards an explanation of Darwin's loss of aesthetic taste. John A. Campbell notes that two explanations have dominated attempts to solve the question of Darwin's affective decline. On the one hand, Campbell suggests that decline has been attributed to Darwin's overconcentration on scientific studies, and, on the other, to his loss of religious faith (Campbell 1974: 159). Two types of evidence, states Campbell, have been used to support these claims: first, the testimony of the *Autobiography*, and second, the limited relevance that Darwin in *The Descent of Man* (1871) accords to emotion and imagination in man's future evolution. In Campbell's view, however, these interpretations of Darwin's affective decline are based on an incorrect or partial reading of Darwin's texts, and are in conflict with explicit evidence in those texts. Although there is some generally accepted evidence, such as the above-mentioned aspects of the *Autobiography* and *The Descent of Man*, Campbell considers that claims regarding Darwin's affective decline are made without sufficient attention to other relevant pieces of evidence.

Campbell bases his theory on a distinction between art and nature. Campbell's argument is that while it is evident that Darwin suffers a substantial loss of interest in art – and indeed, affirms that loss himself in the *Autobiography* – no comparable loss of interest in nature can be detected in Darwin's texts, nor does Darwin claim to suffer such a loss.

> Darwin's affective responsiveness to nature did not undergo a decline at all comparable to his decline of interest in art [...] One need not read far in any of Darwin's works to see that one of the most striking aspects of Darwin's emotional response is his manner of describing the natural world. The language of Darwin's descriptions betrays a relationship with the objects of his study that is personal and affective. In his earliest work his praise of nature is expectedly exuberant (Campbell 1974: 161-162).

To be sure, even if Darwin's manner of observing and describing nature did not become completely unemotional, it did nevertheless change: "[a]s he grew older, Darwin's response, without ceasing to be intense,

became less effervescent and more serene" (Campbell 1974: 163-164). Nevertheless, this merely indicates a change in the character of Darwin's emotional sensibility to nature, not a disappearance of that sensibility.

Campbell next makes an attempt to discover whether Darwin's affective response to nature had religious roots. Comparing Darwin to the natural theologian William Paley, Campbell states that they share "a delight in the particular, but the difference between them is that for Darwin the particulars of nature have very little connection with God" (Campbell 1974: 166). And yet, faced with the Brazilian jungle, Darwin had referred to the existence of something more than the corporeal in a human being. Viewing these types of situations as general and the observation of a flower as something particular, Campbell argues that for Darwin, God was a "God of things in general" (Campbell 1974: 167).

In looking for the source of the delight that Darwin takes in nature, Campbell links Darwin's love of nature with his conception of science, suggesting that the emotional force of the first years of the naturalist's work might almost have eclipsed the scientific rigor of his observations, whereas in his later years the situation was the reverse (ibid.). According to Campbell, however, Darwin achieved a balance between science and emotional delight in nature – a balance rooted in his humanist vision of nature (Campbell 1974: 168). Darwin demonstrates this humanism in not evincing any discomfort with comparing human beings with the most humble organisms imaginable, even expressing admiration for such organisms. This acknowledgment of apparently insignificant organisms is what makes the human being worthy of participating in the organic flow of life. Darwin, therefore, says Campbell, "humanizes knowledge through emotion" (Campbell 1974: 173).

Having established Darwin's continued emotional appreciation of nature, Campbell locates the source of his complete dedication to science in his later years in the serious illness that he suffered and that left him without the energy to pursue other studies, including artistic ones: "Darwin's decline of interest in literature and music was not so much part of a larger hostility to art as a response to a life situation which did not allow him a reserve of emotional energy sufficient for its demands" (Campbell 1974: 173).

In *Charles Darwin, the Anaesthetic Man*, Donald Fleming begins his explanation of Darwin's loss of interest in the aesthetic by describing

the tradition of "dissociation of knowledge and sensibility; fact and affect" that, it is argued, leaves people with only the capacity to know and not the capacity to feel (Fleming 1961: 220). In addition, Fleming focuses on the influence of religion: according to him, it influenced Darwin in the worst possible sense, being an important factor in Darwin's loss of interest in aesthetics. Fleming asks: "Why did Darwin experience this atrophy of the aesthetic instincts?" (Fleming 1961: 225) and locates the key to the puzzle in Darwin's experience with the feeling of the sublime, which links his loss of aesthetic taste to his views on religion.

Fleming states that although Darwin never precisely defines what he means by "the sublime", it is clear that "[t]he sublime was associated by Darwin with an upwelling from the depths of the spirit that appeared to set reason aside and prevail over it" (Fleming 1961: 226). Fleming emphasizes the role of various fundamental pillars like scenic grandeur, religion, and the sublime: "Great art by association with scenic grandeur, scenic grandeur with religion, and all three with the sublime, became part of a single universe of experience" (ibid.). It is not difficult, therefore, to detect in Darwin's texts a connection between the feeling of the sublime, achieved through linking art with nature, and religious feeling.

In Fleming's view, if Darwin came to lose his taste for the aesthetic, that was due to his insistence on distancing himself from the religious (Fleming 1961: 227), an insistence that is clear in his works. Given the connection between religious feeling and the feeling of the sublime, Fleming considers it logical that Darwin's effort to gradually distance himself from religion be accompanied by a symmetrical distancing from the sublime and from everything derived from the aesthetics of landscape. There is, also, a further question. Referring to feelings, Fleming notes: "Intense feeling was undesirable in Darwin's own experience as exacerbating his already keen sensitivities [...] Therein lay a tremendous ambiguity at the very heart of Darwin's position" (p. 229). That ambiguity is embedded in the theory of natural selection, which collides with the search for and broadening of the good in religion because it emphasizes "pain, suffering, frustration, and unfulfillment" so that "[a]ny good that comes *of* it, comes *by* evil" (ibid., italics in original). Fleming describes Darwin's view of this as follows:

> To him [Darwin], a God that dwelt in natural selection would be the worst of all possible Gods. For the proprietor of the universe to have to seek for a mere preponderance of good over evil in the world that he made, which was the best that could be said for any progress attained by natural selection, was monstrous to Darwin's eyes (Fleming 1961: 231).

Darwin, furthermore, finds no solace in the claim that suffering is a means to moral improvement: that claim, for him, only makes it all the more unacceptable to think that a benevolent God should have created millions of animals below humans in the animal scale, and these can obtain no moral improvement from the suffering that is supposed to offer an opportunity for it (ibid.).

One might even state that Darwin's rejection of art is due to the fact that natural selection presupposes the very opposite of nature understood as a work of art. Fleming, however, argues that the source of that rejection is more appropriately located in Darwin's determination not to be an accomplice to evil by accepting the dominion of God (Fleming 1961: 232). If one understands the sublime as the act of observing the relentless power of nature, always seen from a distance, that produces the feeling of human insignificance in the face of natural force, it seems logical to argue that Darwin rejects the possibility of accepting the existence of a God that imposes evil from above:

> The chief lie of lying religion for him was that evil could have been inflicted from on high instead of simply occurring. If, by access to the sublime, he should assent to this lie, his act of charity to mankind for uncovering the harsh necessity of natural selection would fall to the ground (Fleming 1961: 233).

The only possibility of incorporating art into Darwin's system would, therefore, be the discovery of a constructive role for religion in the evolution of humanity:

> The iron band that clamped art, sublimity, and religion together in his own experience would have meant that the obvious way to build art into his system would be to assign a powerful role to religion as a constructive force in the development of mankind (Fleming 1961: 235).

4. Conclusion

A variety of explanations have been offered for Darwin's loss of aesthetic taste. Yet one should not, perhaps, understand this loss as the complete

elimination of aesthetic taste, but rather as a modification produced by the move from an *enchanted* explanation of the workings of nature to an explanation that does not require a mystical element. Darwinian evolution represents one of the most prominent milestones of the process of desacralization or secularization of the world that modern culture entails. Desacralization, however, is not specifically Darwinian as a whole, only in its application to living things and particularly to the human condition.

A scientific naturalism such as Darwinian evolutionism implies a complete conceptual revision of basic ontological presuppositions relating to nature ('species', 'substance', 'natural law', 'teleology', etc.). These modifications have correlations in nature perceptions, including their aesthetic perception and, therefore, also in nature descriptions. Contrary to what is sometimes claimed, the secularized view of the world that Darwinism promotes carries with it neither a devaluation of moral or aesthetic sentiments nor the dehumanization of existence.

Literature

Beer, Gillian. *Darwin's Plots. Evolutionary Narrative in Darwin: George Eliot, and Nineteenth-Century Fiction.* Cambridge: Cambridge University Press, 2009.

Campbell, John Angus. "Nature, Religion and Emotional Response: a Reconsideration of Darwin's Affective Decline." *Victorian Studies* 18, no. 2 (1974): 159-174.

Darwin, Charles. *Journal of Researches into the Natural History and Geology of the Countries Visited during the Voyage round the World of H.M.S. Beagle.* London: John Murray, 1913.

——— *On the Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life.* London: John Murray, 1876.

——— *The Life and letters of Charles Darwin. Including an Autobiographical Chapter,* edited by *Francis Darwin.* London 1887.

Fleming, Donald. "Charles Darwin, the Anaesthetic Man." *Victorian Studies* 4 (1961): 219-236.

Kohn, David. "Darwin's Ambiguity: The Secularization of Biological Meaning." *The British Journal for the History of Science* 22, no. 2 (1989): 215-239.

——— "The Aesthetic Construction of Darwin's Theory." In *The Elusive Synthesis: Aesthetics and Science*, edited by A. I. Tauber, 13-48. The Netherlands: Kluwer Academic Publishers, 1997.

Levine, George. *Darwin Loves You. Natural Selection and the Re-enchantment of the World.* New Jersey: Princeton University Press, 2008.

Paradis, James. "Darwin and Landscape." *Annals of the New York Academy of Sciences* 360 (1981): 85-110.

Sloan, Philip R. "The Sense of Sublimity: Darwin in Nature and Divinity." *Osiris* 2nd Series 16 (2001): 251-269.

Spinoza, Baruch de. *Ethics,* translated by Andrew Boyle and edited with a revised translation by G.H.R. Parkinson. London: Everyman Library 1989.

The Puntel-Whitehead Method for Philosophy

Benjamin Andrae
(Munich School of Philosophy)

The point of departure for this essay is the observation that it is not only metaphysics that is called into question in modern times, but all of philosophy. The basis of this impression, as argued in the first section, is the lack of real progress in philosophy – especially if compared with the rapid advances that most other sciences have made in the 20th and 21st centuries. One main reason for this failure of philosophy is analyzed in the second section: the deductive method that many philosophers apply is found to be inherently flawed. Especially in metaphysical debates, the claim that we can somehow intuit absolutely certain propositions on which to base the whole discussion is fruitless, as an example from the metaphysics of free will shows. Here and elsewhere, differing intuitions leave us with unhappy stalemates which stand in the way of progress. To avoid this situation, an alternative method is presented in the third section: We should not consider the truth as the starting point of a philosophical way of thinking, but rather as its ultimate goal. Instead of starting with intuitively true propositions, we start with vague explananda – as they are presented to us by common sense or by other theoretical frameworks that we already find ourselves in. The goal of philosophical explanation is then to give a unified account of many of these explananda, and thus render them as intelligible as possible. By presenting this method via the thoughts of Alfred North Whitehead and Lorenz Bruno Puntel – who are two of the most important, most revolutionary and most systematic and metaphysical thinkers of the 20th and 21st centuries – it is shown that such a method does not have any inherent limitations, and can reach great speculative depth. We just have to remember that such metaphysics gains its justification from its ability to make all of reality intelligible, rather than from some mysterious insight into the nature of things.

1. Is philosophy 'just empty talk'?

Most, if not all, of us who have made academic philosophy their – temporary or permanent – professional occupation know the colloquial challenge: "All you people do is talk all the time, with nothing ever being decided." Now sometimes, upon our vehement opposition to such a simplistic and non-constructive remark, a clever conversational partner

answers: "Then name one, just *one*, issue that philosophy has settled to universal agreement in the past two-and-a-half millennia."

Now we are in trouble. Really? Are we? Different retorts spring to mind: You might want to simply argue that this is not true, i.e. that there are lots of definite philosophical results.

Another way of answering might be to claim that other disciplines do not fare any better; that huge revolutions of thought have happened and will happen in every discipline, and thus there is no such thing as universal agreement. Thus, the failure to deliver permanent results would not be a special – blameworthy – feature of philosophy.

Both these ways of answering, though, do not usually convince. They both fall prey to what I call the 'time-traveler-objection': Imagine a well-educated person, say someone with a masters degree in some subject, were to travel back in time to talk to the founders of their discipline. In mathematics, physics and biology, in history, geography, medicine and psychology, even in education and economics, the master student would certainly make quite an impression on his historic predecessor. Talking about the advances in modern geometry to Euclid, about modern western medicine to Hippocrates, about the theory of relativity to Newton or even about modern psychological methods and theories to Freud would be quite an event. Probably the new developments would not always be greeted, but certainly the old masters would consider you a scholar of first rate, if not a genius.

On the other hand, talking modern philosophy with Plato (or, even worse, with Socrates) is not – I imagine – something where one would come off as a genius. Rather, I guess, modern theories would be confronted with the same arguments that are given in the classical texts, which are also still put forward in modern debates by experts on ancient philosophy.

And while some of the modern ideas would probably excite the interest of the ancient philosophers, I guess that there would be just as many that they would look upon as thoroughly misguided – and no amount of argument would make them change their minds.

This idea of the time-traveler shows that both ways of answering the above challenge – that there is no real progress in philosophy – do not work. There really *is* a difference. Philosophy does not make progress like other disciplines do.

Maybe the way to answer the critic is to embrace this fact. Again, there are two possible strategies to argue that, while yes, philosophy does not seem to make appreciable progress, it is still a worthwhile enterprise, indeed very much so.

The first strategy is to say that in philosophy, the goal is a different kind of knowledge, and that while this knowledge does not know real progress, it still – and with great effort – needs to be sought for anew by each generation. Kierkegaard formulated this idea by saying: "Whatever one generation learns from another, it can never learn from a predecessor the genuinely human factor. In this respect, every generation begins afresh, has no task other than that of any previous generation, and comes no further, provided the latter didn't shirk its task and deceive itself."[1]

The second strategy to argue in favor of philosophy, in the face of its apparent lack of progress, is to say that *philosophy is very, very difficult*. There *is* progress, this argument claims, it is just much, much slower than in other sciences. This difficulty, the argument continues, stems from the core of the philosophical enterprise itself: It is at the end of the food-chain of explanation. Peter Simons puts this as follows: "One of the problems about being a philosopher is that you cannot shove the hard questions off onto someone else. Other sciences can leave a hard question aside, saying, 'That's a problem for philosophers, not for us.' The buck stops with us."[2]

Further, and partly as a consequence of having to answer the left-over questions from every other discipline, philosophy has to, and wants to – in a way – *be about everything.* It has a fundamentally universal aspiration, attested to here by Jorge Gracia: "Only philosophy aims to be fully comprehensive. Philosophy aims to produce a big picture, even of partial aspects of the world; it is not content with partial pictures of the world, or any of its parts."[3] These two – one could say 'metaphysical' –

---

[1] Søren Kierkegaard, *Fear and Trembling* (London: Penguin Books 2003), 145 (in the epilogue).

[2] Peter Simons, *Interview in Metaphysics: 5 Questions,* ed. Asbjørn Steglich-Peterson (Copenhagen: Automatic Press / VIP 2010), 138.

[3] Jorge Gracia, *Metaphysics and its Task* (Albany: State University of New York Press 1999), 11.

aspects certainly makes philosophy much harder than any other discipline.

What are we to make of this second pair of arguments? They both contain a meaningful insight, and, unlike the previous arguments, cannot be discarded as simply wrong. Yet, they both do not quite satisfy. The first, essentially existentialist, argument shows that philosophy – as it is sometimes practiced – has a non-theoretic aspect. Some philosophers are not simply after a true theory, but after much more. That *more* might be called wisdom, spiritual insight, moral uprightness, beauty or even salvation. Kierkegaard, in the paragraph following the above citation, says that the 'humane factor' he is speaking of is 'passion', is 'loving'.

I do not wish to belittle any of these goals of philosophers – which are absolutely worth pursuing – but one cannot help but notice that their attainment is not really promoted by what academic philosophy does – which is publish articles and books containing theoretic sentences.

As long as it is granted, as I think it should be, that there is a meaningful way to use the term 'the discipline philosophy', and that this term denotes a fundamentally *theoretical* activity – engaged in, during at least the last few centuries, by women and men in universities, just like any other theoretical activity – let me restrict the above arguments to this use of the term. The kind of knowledge sought after in this discipline, a true theory, is not so fundamentally different from the knowledge that other disciplines seek.[4] Consequently, the problem that this theoretical endeavor does not seem to make a lot of progress appears with undiminished force, and the Kierkegaardian way of thinking does not apply anymore.

The second argument, that philosophy is just *so very hard*, due to its

[4] Of course, this academic discipline of philosophy can – as it in fact does – include *theories* of ethics, beauty, salvation, etc. But a theory of beauty wants to be true, not beautiful, and a theory of salvation wants to be true, and not bring about salvation in either the reader or the writer. Note also that theories can be integrated in other endeavors, as when someone utters the normative (and thus non-theoretic) sentence: "Find out what is the right thing to do and then do it!" But such a sentence can never be a part of academic philosophy, conceived as (merely) the search for a true theory.

fundamental and *universal* nature, also sees something that is correct – namely precisely this dual character that sets philosophy apart from any other academic discipline. And yes, it is thanks to this character that philosophy is more difficult than other endeavors: in philosophy, and especially in the part of philosophy that is called 'metaphysics', we tend to go all the way and offer a final interpretation of matters, *and it is precisely this that interests us in the first place.* Certainly, this is very hard.

The second argument goes quite a long way in rationalizing the apparent lack of real progress in philosophy.

Yet, as an excuse, saying that the progress is ever-so-slow because the task is so very hard comes off as a bit feeble. More importantly, it somewhat misrepresents how academic philosophy fails to make a lot of progress – which is precisely by *neglecting* its universal aspiration, and by misunderstanding its fundamental one. It seems that for every philosophical position these days there is at least one that claims the exact opposite, and very little effort or hope to reconcile the two. As for the fact that philosophy wants to – and has to – offer final interpretations, the problem is that these are understood as 'the basis', or the foundation, of thought. Therefore, whenever there is change to these 'basics', the whole building of philosophy collapses (for those who support the change) and a new one is erected. This leads to numerous positions which are not just opposite to a given position, but which are – or claim to be – utterly incomparable to it, giving rise to the impression that nothing is ever decided in philosophy.

So, while the *theoretical, fundamental and universal science* that is academic philosophy is certainly very difficult, this difficulty is not the only culprit for its lack of progress, and cannot completely account for it.

Rather, we are faced with a situation in which we have to agree, only to a certain extent and after important clarifications, with the critic: academic philosophy really does not always make enough definite progress, and thus becomes vulnerable to the charge of being 'just empty talk'[5].

[5] After I had written this paper, it was brought to my attention that David Chalmers has an unpublished paper on his website (at http://consc.net/papers/progress.pdf) called, 'On the

However, in analyzing the attempts to counter the critic, we have made important steps towards diagnosing just what the problem might be.

2. Starting out with the truth

Most philosophical arguments come down to a *modus ponens*: One affirms statement A, and the fact that A implies B, and from that concludes that B is true. An example, taken from the metaphysics of free will, would be: "Given the truth of Determinism, and the fact that Determinism does imply that there is no free will, it follows that there is no free will." Obviously, there is nothing wrong with this logically. In standard propositional logic $((A \wedge (A \rightarrow B)) \rightarrow B)$ is a theorem. More importantly, this is the way we naturally deduce a truth from other truths: "It is raining outside, and I know that *if* it is raining outside, *then* I do not need to water the plants in the garden today; so, I do not need to water the plants in the garden today." No one in his right mind would – on philosophical or logical grounds – want to criticize this line of thought.

How, then, does this tool fare in philosophical discussion? Coming back to the metaphysics of free will, there is a lot of debate about the conclusion, namely that there is no free will. In fact, this is a prime example where no definite result of any kind is in sight. What do these philosophers debate? And how is this different from the case of the rain in the garden?

The difference lies in the fact that in the philosophical argument the premises are far from certainly true. In the case of the rain in the garden, one usually relies on one's senses to affirm the fact that it is raining, and on one's theoretical and practical experience with botany to affirm the fact that with enough rain plants do not need extra water.

Limits of Philosophical Progress', where he shares my diagnosis that philosophy has not made enough progress on the big questions, and says about this fact: "To me, this is the largest disappointment in the practice of philosophy." (p. 4). He discusses the possibility that the method of philosophy is partially to blame – as I will claim below – but does not offer an alternative – as I *will* do below.

Neither of these inputs is subject to reasonable doubt, under normal circumstances.

In the philosophical debate, on the other hand, the truth of 'Determinism obtains' and/or 'Determinism does imply that there is no free will' is very much in question. So, a debate trying to decide the truth of one statement – 'there is no free will' – has turned into a debate trying to decide the truth of (at least) two other statements. So far, still, there is absolutely nothing wrong with this: the truth-values of these three different statements are definitely somehow entangled, and thus one cannot decide on the truth of one of them without regarding that of the others.

Yet, factually, the situation has devolved somewhat. What we are stuck with is an argument about the logical form of which there is no question, together with different people assuming different truth-values for the premises of the argument, obviously leading to different conclusions.[6] This situation repeats itself in quite a few other philosophical debates, leading to unhappy stalemates and certainly not to any progress in the project of obtaining a universal and fundamental theory.[7]

So, how can we go about breaking such stalemates? Here it becomes problematic. If we stick to the above method of arguing, we need another *modus ponens* that has the problematic premises in question as conclusions. But of course, this second *modus ponens* would again have premises. The process would have to be repeated until we hit premises that are self-evident, or at least practically indubitable, like in the 'rain in the garden' case. Unfortunately, in many cases no such deductions from universally self-evident truths can be given for a philosophical debate, and all one ends up with is different 'intuitions' of what is self-evident

---

[6] This can lead to both sides accusing the other of 'begging the question'. An example concerning the free-will debate, where the difficulties involved in this accusation come out nicely, is: John Martin Fischer and Garret Pendergraft, "Does the consequence argument beg the question?", *forthcoming*.

[7] It almost seems as if modern philosophers have sussed out Socrates' trick of making people agree to seemingly harmless premises and then deducing uncomfortable conclusions – to such an extent that nowadays many philosophical debates very quickly lead to: 'wait wait wait, I do not assume this premise'. *If you don't like the conclusion, simply deny one of the premises* appears to be the motto.

for different philosophers – rather than real progress.

In the face of this dilemma, one might be tempted to go for what is sometimes termed 'conditional philosophy' – i.e. arguing only for certain inferences of the form 'if you assume this, then it follows that' – without affirming the first premise. More generally, this amounts to uncovering ways in which philosophical questions are related, without ever asking which actually is the correct theory. While such work has considerable merit – it is usually very difficult – this leaves something to be desired if we actually want definite progress.[8]

No. If we want such progress, we have to turn to altogether different methods. In what follows, I will go on to showcase one such method, based on the works of Lorenz Bruno Puntel and Alfred North Whitehead – who are two of the very few systematic and metaphysical thinkers of modernity – which is based on the idea that *truth is not the starting point of philosophical thought, but its final stage*. Thus, as I hope to show in the next section, it avoids the unhappy stalemates described above and might have the potential of resulting in some real philosophical progress.

3. The Puntel-Whitehead Method

Let me start with a few words of warning: Puntel's and Whitehead's ideas about methodology are deeply intertwined with the rest of their philosophical systems, and thus a greater understanding of those ideas can be gained by studying their respective works *in toto.*[9] Since I cannot provide such a grand perspective here, we will have to make do with a look at the main points, densely focused on dealing with the issues

[8] This 'conditional philosophy' is what I think many of the publications about the metaphysics of free will are about these days.

[9] The main sources for these works are: Lorenz Bruno Puntel *Structure and Being* (University Park: The Pennsylvania State University Press 2008) – in the following abbreviated as SB – and: Alfred North Whitehead, *Process and Reality.* Corrected Edition, ed. David R. Griffin and Donald W. Sherburn (New York: Free Press 1979). Originally published in 1929 – in the following abbreviated as PR. Both abbreviations are then followed by the page number.

discussed in the first two sections.[10] Let us get started with that:

Puntel, in one sentence, sums up the difference between his method (the 'systematic' one, which is yet to be explained) and the 'foundationalist' one (described in the last section) as:

> "The foundationalist process is, to put it one way, 'archaeologically' oriented, whereas for the systematic process of grounding, the designation, 'teleological' is appropriate. [...] a systematic philosophy is metasystematically grounded, not when it is foundationally established, that is, not when it is developed on the basis of foundational theses that are established as they are presented, but only when, after it has been completed, it is considered as a whole."(SB 68)

In a very similar vein, Whitehead says:

> "Philosophy has been haunted by the unfortunate notion that its method is to indicate dogmatically premises which are respectively clear, distinct, and certain; and to erect upon those premises a deductive system of thought. But the accurate expression of the final generalities is the goal of discussion and not its origin. [...] The verification of a rationalistic scheme is to be sought in its general success, and not in the peculiar certainty, or initial clarity, of its first principles."(PR 8)

The clarification of both these centrally important statements of the basic idea about how progress in philosophy can be achieved is the goal of most of the rest of this essay.

The first issue to be clarified will be this: If we cannot start with affirmed premises, then where *do we start*? Despite its highly theoretical nature, this is also a very pragmatic question. I want to have a good philosophical theory about this-and-that, how do I go about obtaining one? What should I do first?

The answer is that precisely the above mentioned 'this-and-that' is your starting point. The things that you want your theory to explain, with all their vagueness and indeterminacies, with their inherent demand to be clarified, are the points of departure.

A positive aspect – again from a pragmatic standpoint – is that here the chance for agreement between different philosophers and even

10 As a consequence, the differences – which are definitely present even at the methodological level, and much more so when looking at each of their respective systems as a whole – between the ideas of Puntel and Whitehead are not the main concern of this essay.

schools of philosophy is much higher. We do not have to agree to the final, determinate form of the theory, but we have to agree what it is that we actually want to explain.

I will call this starting point the initial 'datum', in the following sense:

> "Our datum is the actual world, including ourselves; and this actual world spreads itself for observation in the guise of the topic of our immediate experience. The elucidation of immediate experience is the sole justification for any thought; and the starting point [...]. But we are not conscious of any clear-cut complete analysis of immediate experience." (PR 4)

> "The dimension of the universe of discourse is the comprehensive datum in the literal sense: what is given to philosophy to be conceptualized and/or explained (i.e., everything with which philosophical theorization can and must be concerned)." (SB 10)

It is important to note that this sense of 'datum' cannot be restricted to any of the usual senses of the word; neither are we restricted to positions of pointers in what we want to explain, nor in the impressions of our senses – although these things certainly are among the data. The term simply includes everything of interest: the price of a cup of coffee at my favorite cafe, the color of the Mediterranean on a cloudy morning, the various demands of my conscience to act in one way or another, the nature of theoretical implication itself and its various possible formulations, the mathematical structure of general relativity, last night's dreams, Darwin's theory of evolution or the feelings evoked by a particular poem on a particular day are but a few examples of such data.

This wide conception of data is, in my view, one of the most important aspects of this method. It means that philosophy – thus conceived – is at the core a positive, productive effort that cannot and should not exclude any facets of the world as mere appearances or illusions, or as falling categorically outside of its scope. Whitehead says: "Philosophy may not neglect the multifariousness of the world."(PR 338)

Yet, we must always remember this: data are not 'true', for truth will have to wait until the *end* of inquiry. They are merely the explananda, the things that could and ultimately should be explained by a theory. Puntel (following Rescher) says: "'Datum' here can be understood as a candidate for inclusion in a theory or for truth." (SB 11)

It is worth noting that, since we do not build our philosophy upon

as-evident-as-possible assumptions, there is no privileged starting topic for philosophy according to this method; nothing needs to be settled first, before other questions can be approached. We can start wherever we want, and thus the very many different projects engaged in by modern philosophy are all prima facie relevant. The only thing one must keep in mind is the universal aspiration of philosophy, which is to eventually include and *integrate* (in a sense to be detailed below) all the data.

Now, one crucial question needs to be answered regarding these 'data': Are they somehow pure, as in 'independent of every theory'? The answer comes from Puntel, and is a resounding 'No!' – and it already includes directions for how we are to proceed from the data:

> "The datum, thus the particular data, is/are available as prestructured, at the fundamental or zero-level of theorization, within everyday theoretical frameworks relying on ordinary language, and on higher levels of theorization within the theoretical frameworks of the various sciences, including philosophy. They include all the 'somethings' that emerge as articulated theoretically in the universe of ordinary discourse when there is talk of 'things', 'the world', 'the universe', etc. Systematic philosophy must attend to these and to relevant higher-level articulations and attempt to bring all these data into a comprehensive theory. Doing so does not involve accepting such data as in any important sense "ready-made" components of the theory; quite on the contrary, they are precisely candidates for restructuration within the theory, items that must be conceptualized and explained, a process that involves radical corrections and transformations." (SB 11)

Here we see two centrally important points: (a) The fact that 'data' are always already within what Puntel calls a 'theoretical framework'. Thus, if I want – coming back to the previous example – a good theory of the free-will issue, I will start with 'Determinism within mathematical physics' and/or 'Determinism as divine foreknowledge in theology', and with 'Freedom as concept in everyday thought' and/or 'Freedom in moral theory' and/or 'Freedom in the theory of law', etc. Also, we see that (b) the suggested method does not – not *at all* – content itself with collecting different facts, e.g. uses in ordinary language, and ordering them a bit. On the contrary, it does not shy away from radical re-interpretation (Whitehead and Puntel offer two of the most innovative metaphysical systems available!). The method starts with what we have, and proceeds boldly from there. How, exactly, it does that, will be the next issue for our discussion.

Once we have our data, the first thing we have to do is articulate them. For example, when looking at 'Freedom as a concept in everyday thought', we will look at the different parts of the notion *within the realm of everyday thought* and their connections. Preferably, this articulation takes some simple logical form, such as propositional logic. Here, philosophers need to practice restraint, and tolerate what – from our point of view – must look horribly vague at times. Note that this has the practical advantage that agreement is very likely; no-one's *philosophical* theory is at stake yet, all we do is lay out what it is that we want to explain. Of course, if the datum stems from a more theoretical area, e.g. mathematical physics, good knowledge of the datum's position within that area – and a more fine-grained structure than propositional logic – is required.[11] Often this very simple first step of articulating what it is that one is actually trying to explain is a decisive factor in weeding out misunderstandings, and thus should be carefully undertaken.

Next, we take our datum, which is always – as we have seen – already located within some theoretical framework, and collide it with another datum. In the free-will debate, for example, we collide the datum of 'Determinism' with that of 'Freedom'. But since they are both set in their respective theoretical surroundings, such a 'collision' is only possible if we formulate a new theory that includes both the data. We have to *speculate*, and try to come up with a theory that *includes,* but *transforms* the initial data. Only if the new theory can include both data, have we been successful, and we have a better theory than we had before – and thus we have made real philosophical progress. Whitehead describes this in his famous metaphor:

> "The true method of discovery is like the flight of an airplane. It starts from the ground of particular observation; it makes a flight in the thin air of imaginative

[11] It is possible, though, that at this stage some theory proves inconsistent. In this case, we either abandon the attempt at explanation of all of it and restrict our attention to some consistent 'subset' of the data, or we 'live with' the inconsistency, hoping that it will disappear along the way to more comprehensive theories. Further note that circularity, of the form $A \rightarrow B \rightarrow C \rightarrow A$, or partial circularity, should not be considered a lethal flaw for a theory at this stage. All we do is articulate the structure of a datum, and if it includes such loops, then so be it. We are – after all – not looking for self-evident premises to affirm!

> generalization; and it again lands for renewed observation rendered acute by rational interpretation." (PR 5)

How are we to understand that a theory 'includes, but transforms' the data? Something has to remain the same about them and something has to change about them. To put it briefly: What remains the same is the structure of the initial data, what changes is their position regarding the rest of the (now much bigger) system. An example could be how, even after integration into a theory of 'Human agency and natural laws', the datum 'Determinism within mathematical physics' still maintains its mathematical shape; for example the Picard–Lindelöf theorem still holds, and relevant differential equations still have unique solutions. What changes is how we interpret this mathematical structure, that is, how the different elements in it relate to the other parts of the theory. This is what Whitehead means when he says that 'renewed observation is rendered acute by rational interpretation'.

Note that this interpretative inclusion in a wider framework also can, and should, provide many conceptual relations between different data that were previously absent: Now, within the larger theory, we gain a better understanding of the terms than we previously had within our smaller theories, as they are also explained in terms of one another. Coherence – *understood as intelligibility* – has increased.

Possibly, and indeed probably, not all the vagueness that we felt lay in the initial data, has disappeared yet. Then, what we do is, we include more data. If – for example – you are not satisfied with your theory of 'Human agency and natural laws', even after formulating it in a systematic way that includes all the initial data, then drag in data from the philosophy of mind.

In fact, according to the universal aspiration of philosophy – which strives to be not just a conjunction of different topics, but a unified theory – any datum that has not yet been included will keep 'pushing in', asking for expression in the theory, forcing us, again and again, to give up the orthodoxy of a closed theory to admit the next datum, leading, in turn, to a just-a-bit-more-intelligible theory. This openness to data, in contrast with a tendency to explain away or ignore anything that does not fit a theory, is something that Whitehead thought very important, and admired about Locke: "The enduring importance of Locke's work comes

from the candor, clarity and adequacy with which he states the evidence. [...] He explained, in the sense of stating plainly, and not in the more usual sense of 'explaining away'" (PR 145).

Eventually – this is the hope – we will end up with a "coherent, logical, necessary system of general ideas in terms of which every element of our experience can be interpreted" (PR 3). Of course, there are problems with this notion of an ideal system: Can we really ever get there *and know we are there*? Is it *certain* that there is just one of these? Is it *always clear* which theory includes more data, or is more coherent? It seems to me that the answer to all these questions might well be 'no'.

Yet, the mere fact that we can ask them shows that we are capable of quite general thought, so there is hope for achieving relatively universal systems of thought that respect hugely diverse data and integrate them coherently. And every attempt along this road is surely better and more constructive than the 'unhappy stalemates' between different positions – each denying the other's assumptions – which I described above. Therefore, clearly, the Puntel-Whitehead method shows much more promise of leading to real progress in academic philosophy.[12]

4. Other coherentist approaches and the inclusion of natural science

Of course Puntel and Whitehead are not the only ones who came up with the idea that the search for knowledge in philosophy should proceed in a broadly coherentist kind-of-way. Puntel quotes Rescher

[12] What about the truth? Did I not promise that, with this method, the truth lies, if not at the beginning, then at least at the end of the investigation? Well, in order to get one's hands on the truth, one needs a philosophical system capable of expressing what 'truth' is. And this is how it should be. For what is truth other than a theoretical notion to be decided within a theoretical framework of some sort? The framework will have to be very universal, and will have to include metaphysics in a very robust sense. Lack of space prohibits the presentation here, but let me mention that both Whitehead and Puntel (the latter in much more detail) develop a theory of truth within their attempts at a comprehensive philosophical system. Using a strong notion of truth at any previous step within the methodological process would – and in fact often does – lead only to different parties contradicting each other over who is 'right', without a clear notion of what 'being right' really means. This inner-theoretic notion of truth also disarms the problem faced by coherentist approaches that define truth extra-theoretically as maximum coherence, which is that then, since we probably never reach the ideal system, we never say anything true.

extensively, and Whitehead certainly read Bradley. So, to a certain extent, the choice of these two philosophers as champions of this method is arbitrary and reflects personal preference. But that is not the whole story. Whitehead and Puntel have also distinguished themselves through the fact that they have tried to live up to the challenge of the above described method all the way and have both formulated a universal philosophical system.[13]

We see, thus, that this method does not have some inherent limitations – as coherentism is sometimes thought to have[14] – but is, in fact, capable of fulfilling the demand that philosophy is a universal and fundamental discipline, stopping with nothing short of an explanation of everything.

In the modern context, it is especially important that both these systems *include* natural sciences, but reach metaphysical conclusions that are in no way reducible to them. This makes them viable in a world where the existence and importance of natural sciences is hardly deniable, but where their status as a metaphysics-against-their-will has brought us nothing but problems. We see, thus, that the method suggested here does not go against the scientific method – in fact it accepts all experiments and scientific theories as perfectly valid and highly interesting data; what it does is, it notes that there are many things that interest us, that need explaining (i.e. are data in the present sense), which are not in any way treated by the natural sciences. Persons, ethics, modality, time, emotions, mathematics, art, religion, history, etc. are a few examples.[15]

---

[13] What these systems are like and whether they wholly adhere to the methodology described cannot be discussed here. Also, the fact that these systems are very different in a fundamental aspect – Whitehead is concerned with 'taking mental experience *really* seriously' and Puntel with language as 'taking the linguistic turn *really* seriously' – could be seen as a serious problem for the method described here; for it seems that there is no real progress now anyway. Although this question is extremely difficult to answer (it involves two whole metaphysical systems!), I think that there are points of contact that show there is a possibility of comparison and integration of some sort between to two (one of them could be a 'language of thought', the possibility of which Puntel mentions briefly in a centrally important passage in SB 96).

[14] Puntel characterizes Rescher's position as "fundamentally epistemically rather than ontologically oriented" (SB 48), and considers that a shortcoming.

[15] Simply saying that all of these things are reducible to natural science is, by the way, a

That a wider conception of data is called for, one that includes more than positions of pointers, has been noticed by other philosophers – although it has not been precisely articulated: Kriegel compares scientific data to "the role of data being played by pre-theoretic verdicts about what falls in the extension of a certain phenomenon"[16] in philosophy, and thinks that '(quasi-)empirical adequacy'[17] to these – more widely conceived – data is one of the virtues a theory should have. Chisholm seems to think along the same lines when he says that certain facts, e.g. about persons, are "pre-analytic or pre-philosophical data. Any philosophical theory which is inconsistent with any of these data is *prima facie* suspect."[18]

Including such data, along with ones from the natural sciences – but neither data as 'ready-made', but merely as the starting point of our philosophical theorizing – and bringing them together into one universal, or at least rather general, way of thinking, is one of the many things that the Puntel-Whitehead method, as I have called it, is good at and good for. The question most people – including scientists from other, possibly neighboring, fields – ask themselves upon hearing a specific result of natural science is: "What does that *mean*?" The answer to such a question cannot be anything else than an integration of the scientific result with other data, or other theories, that the person asking has about herself and the world. If this second set of data were not included, the spirit of the question "What does that mean?" would not be answered. The result of the integration, using the method suggested here, would be a new theory that adequately includes all the data in a coherent way, but goes beyond all the component theories in central – possibly metaphysical – ways. Such a theory then increases the intelligibility of the world.

And, I dare say, if a time-traveler could explain to Plato how his ideas are now integrated in a much wider unified theory, including the philosophy of language, of mind, of modality, including different natural

---

philosophical thesis – and not a scientific one – and thus requires philosophical argumentation.

[16] Uriah Kriegel, *The Sources of Intentionality* (Oxford: Oxford University Press 2011), 172.

[17] Kriegel, *The Sources of Intentionality,* 172.

[18] Roderick M. Chisholm, *Person and Object* (London: George Allen & Unwin 1976), 18.

sciences, etc. – and all the important ideas and arguments these and other fields have developed over the years – *he would be impressed.*[19]

Literature

Chisholm, Roderick M. *Person and Object.* London: George Allen & Unwin, 1976.

Fischer, John Martin & Pendergraft, Garret. "Does the consequence argument beg the question?" *Philosophical Studies*, forthcoming.

Gracia, Jorge J. E. *Metaphysics and its Task.* Albany: State University of New York Press, 1999.

Kierkegaard, Søren. *Fear and Trembling*. London: Penguin Books, 2003. Translated by Alastair Hannay. Originally published in 1843.

Kriegel, Uriah. *The Sources of Intentionality.* Oxford: Oxford University Press, 2011.

Puntel, Lorenz Bruno. *Structure and Being: A Theoretical Framework for a Systematic Philosophy.* University Park: The Pennsylvania State University Press, 2008.

Steglich-Peterson, Asbjørn. *Metaphysics. 5 Questions.* Copenhagen: Automatic Press / VIP, 2010.

Whitehead, Alfred North. *Process and Reality: An essay in cosmology.* Corrected Edition, edited by David R. Griffin and Donald W. Sherburn. New York: Free Press, 1979. Originally published in 1929.

[19] At least by our theoretical achievements, not necessarily by our moral and/or spiritual ones.

On Rudolf Carnap's *Aufbau* and its Interpretations

Rafał Wodzisz and Piotr Lipski
(John Paul II Catholic University of Lublin)

In this paper we deal with a subject which gained its importance in the 1990's. At that time, scholars, such as Alan Richardson and Michael Friedman, raised the question about how Rudolf Carnap's *Der logische Aufbau der Welt* should, or can, be interpreted. More importantly, they raised several objections against the received view of Carnap's early work and pointed to a wider variety of inspirations that influenced him. The received view was established mainly by Wilard V. O. Quine, Alfred J. Ayer and Nelson Goodman. All three of them read *Aufbau* as an attempt to realize the project of empiricists formulated at the beginning of the 20th century by Bertrand Russell. That is the basis of the phenomenalist, reductionist reading of the discussed book. This view focuses on the particular constitutional system developed by Carnap in *Aufbau*. Friedman and Richardson, on the other hand, argue quite convincingly that Carnap's aim was more general and consisted of creating a constitutional theory, a methodology of creating all sorts of systems, and that phenomenalist system presented in the book serves only as an exemplification.

1. Introduction

Although Rudolf Carnap is one of the most prominent philosophers of the 20th century, his accomplishments were often neglected in the past. It was, to some extent, due to his association with the radical positivists of the Vienna Circle, who were very anti-metaphysical in their approach. Carnap's work was straightforwardly interpreted as positivistic and anti-metaphysical. This interpretation was mistaken, especially with respect to his early views, developed in *Der logische Aufbau der Welt*.[1] In this paper, we wish to investigate how the received view of *Aufbau* came about and how some recent investigations contributed to the change that has occurred with respect to the reception of the book.

[1] Later on referred to as *Aufbau* for convenience.

We carry out our task in the following order: Firstly, we will present the project of *Aufbau,* and it will be done by summarizing the book. Secondly, we will give an account of classical interpretations of *Aufbau*. Thirdly, modern interpretations of the given book will be described. Finally, we will attempt to discuss a change that occurred with respect to the reception of *Aufbau*, i.e. a shift from classical to modern interpretations. To achieve our goal, we take a mostly historical approach. We will analyze the works of the authors who devoted a significant part of their professional careers to deal with Carnap's *Aufbau*.

Throughout the text we use the Chicago Manual of Style for references, with only one exception: references to *Aufbau* are given merely by indicating the paragraph.

2. The project behind *Aufbau*

Aufbau is divided into five parts, each of which is divided into some chapters, which are in turn divided into paragraphs – the smallest units of the text. Each unit is at most a few pages long. In the first part, Carnap indicates the establishment of a constitutional[2] system as the aim of the whole *Aufbau* project (§1). A constitutional system should be understood as a family tree of all concepts having cognitive content. Some basic terms are 'roots' of that 'tree', and for every single concept it is possible to present its constitution, starting with those basic terms. To constitute some object out of some other objects means to deliver a method of translating every sentence containing a name of the former object into a sentence with names of the latter objects only (§2). [3] What is important for us is the fact that the aim of the *Aufbau* does not lie in any particular form of the constitutional system, but in the mere possibility of constructing any such system. Therefore *Aufbau* should be primarily seen as an exposition of the constitutional theory, i.e. a theory

[2] In Rolf George's translation of *Aufbau,* the term "constructional" is used instead of "constitutional". Throughout our paper we use the latter, which is in accordance with contemporary literature.

[3] Plus the expressions of logic.

presenting formal and methodological aspects of constitutional systems in general.

The second part starts with a very important distinction between two possible kinds of description of any given domain – property description and relation description (§10). The former consists in listing features of members belonging to the given domain, the latter in describing relationships connecting these members. Giving the age of individuals from some group can serve as an example of a property description of the group, whereas indicating the relation of being older/younger, which holds between these individuals, delivers a relation description of the group.

There is a special type of relation description, namely structure description (§12). In order to present a structure description based on some relationship, one has to list all pairs of objects which are connected by that relationship. Similary one can show a diagram of that relationship (where arrows indicate which objects are connected and in which direction they are connected). Yet another way of expressing it is that structure description of any domain means presenting the structure of a relationship which holds between the members of the domain. Thus it should be clear that a structure description delivers only information about formal features of the relationship. Carnap uses structure description to introduce the idea of purely structural definite descriptions (§§13, 14). In general there are two ways of referring to any object, either by pointing at the object (ostensively) or by describing it definitely, i.e. in a manner which allows one to distinguish that particular object without confusion. Structure description of a given domain allows (at least in some cases) one to refer to every member of that domain by description, appealing only to structural features of a relation constituting basis for the structure description (§§15, 16).

In literature there is a well known example of such purely structural definite description, which is due to Alan Richardson[4], who gave an easier version of Carnap's original example (§14). Let us think of a simple railway network consisting of just two lines: one going from

[4] Alan Richardson, *Carnap's Construction of the World. The Aufbau and the Emergence of Logical Empiricism* (New York: Cambridge University Press 1998), 47-51.

West to East, and one leading from South to North. At the crossing point of the two lines there is a central station. Imagine further that there is a different number of stations in every direction from the central station (e.g. one station northwards from central station, two stations southwards, three westwards and four eastwards). The structure of the described railway network may serve as a basis for structure description, which in turn can deliver structurally definite descriptions. The central station, for instance, can be purely structurally and definitely described as the only station which neighbors on four different stations and – as another example – the most northern station can be described in a similar manner, as the only station which neighbors on only one station, and that is the central station. In a similar way, every single station from the above example can be referred to. If some of the structure descriptions are not sufficient to deliver purely structural definite descriptions of every member of the given domain, then one should look for another structure description of that domain in order to obtain those definite descriptions. In principle, it should be possible to refer to every object just by a purely structural definite description.

One remark is required at this point. The above enclosed presentation is highly detailed because it concerns topics which are – as we will see – crucial for understanding interpretations of *Aufbau*. The following summary of Carnap's work will not be so thorough.

The second part of *Aubau* finishes with a distinction between three basic kinds of objects (physical, psychological and cultural) (§§18, 23) and five types of relations holding between objects of these kinds: (psychophysical, expression, designation, manifestation and documentation relations) (§§19, 24). It is quite important because every constitutional system should have, as an outcome, a constitution of objects of all those kinds.

The third part of the discussed book is the longest one and covers constitutional theory. As we said above, constitutional theory is devoted to giving a general exposition of formal aspects of constitutional systems. Carnap divided this exposition into four chapters, each of which presents different kind of forms. Firstly, ascension forms are under scrutiny. These are forms that appear repeatedly throughout every constitution. They can be thought of as bricks used to build more

complex units). There are two such forms: classes and relations (§§33, 34, and 40). The second chapter deals with system forms, i.e. possible shapes of a constitutional system as a whole. It is precisely in this place where Carnap clearly states that the constitutional system with auto-psychological objects as a basis is not a necessity (§54), for it is plausible to obtain a system with (for instance) physical objects as basic ones (§59). The issue of the basic forms is discussed in the third chapter, where the basic elements and the basic relationship for a particular constitutional system, as presented in *Aufbau,* have been chosen (§§64, 78). Finally, there is an overview of object forms, i.e. ways of constituting some particularly important types of objects.

The fourth part of the book contains an outline (in some initial parts very precise and formal) of a specific constitutional system put forward by Carnap in *Aufbau*. The order of constitutions in that system is following:

1. Recognition of similarity is selected as a basic relation of a system (§108) and is, in fact, its only basic concept.
2. Upon this relation all auto-psychological (private, subjective) objects are constituted.
3. Physical objects are obtained.
4. Objects from 3. And 4. serve as a basis for the constitution of psychological objects of other individuals (i.e. hetero-psychological objects).
5. The final step of creating the system exhausts the constitution of the intersubjective world (§148) and cultural objects.

In the last part of *Aufbau* – which is quite typical for Carnap – previously established results are applied to classical philosophical issues. Problems of psychophysical relationships, of reality and of the limits of science are addressed among others.

3. Classical interpretations of *Aufbau*

The popularity of the interpretations discussed below was possible thanks to a delay in the publication of the English translation of *Aufbau*.

It was first published in 1967, and so it was almost 40 years after the original version appeared (1928). During this time, the classical interpretation of *Aufbau* had been established. It is a generally accepted fact[5] that we owe this interpretation to Alfred J. Ayer[6], Wilard V. O. Quine[7] and Nelson Goodman[8]. They read *Aufbau* in the light of logical positivism, and therefore their reading puts the emphasis on the empiricist program that, according to them, can be found in the book. These authors propagated what is now called the 'received view' of *Aufbau*, and this is placed under scrutiny below. The view focused around the main doctrines of the Vienna school: verificationism and radical empiricism, which both suggest a phenomenalist interpretation of *Aufbau*.

Quine describes Carnap's work with the following words:

> Radical reductionism, conceived now with statements as units, sets itself the task of specifying a sense-datum language and showing how to translate the rest of significant discourse, statement by statement, into it. Carnap embarked on this project in Aufbau. [...] In a series of constructions, in which he exploits the resources of modern logic with much ingenuity, Carnap succeeds in defining a wide array of important additional sensory concepts which, but for his constructions, one would not have dreamed were definable on so slender a basis. He was the first empiricist who, not content with asserting the reducibility of science to terms of immediate experience, took serious steps toward carrying out the reduction.[9]

According to Quine, the basic program of *Aufbau* consists of two steps in which a radical empiricist project is being undertaken. Firstly, statements referring to our sense experience have been chosen as a basis

---

[5] Cf. Michael Friedman, *Reconsidering Logical Positivism*, (New York: Cambridge University Press 1999); Michael Friedman and Richard Creath (Eds.), *The Cambridge Companion to Carnap*, (New York: Cambridge University Press 2008); Alan Richardson, "How not to Russell Carnap's Aufbau," *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1990, (1990): 3-14; Idem., "Logical Idealism and Carnap's Construction of the World," *Synthese* 93, (1992): 59-92; Idem., *Carnap's Construction of the World*.

[6] Alfred J Ayer, *Language, Truth and Logic*, (New York: Penguin Books 1978).

[7] Willard V. O. Quine, *From a Logical Point of View*, (New York: Harper Torchbooks 1963); Idem., *Ontological Relativity and Other Essays*, (New York: Columbia University Press 1969).

[8] Nelson Goodman, *The Structure of Appearance*, (Cambridge: Harvard University Press 1977).

[9] Willard V. O Quine, *From a Logical Point of View*, 39.

– phenomenalism. Secondly, by means of modern logic, the complete body of our knowledge was supposed to be translated into those basic experiences – reductionism. This idea has been tied up by Quine with Bertrand Russell's empiricist program:

> [...] in *Our Knowledge of the External World* and elsewhere he addressed himself to the epistemology of natural knowledge, on its conceptual side.
>
> To account for the external world as a logical construct of sense data – such, in Russell's terms, was the program. It was Carnap, in his *Der logische Aufbau der Welt* 1928, who came nearest to executing it.[10]

Quine read *Aufbau* as the latest attempt to realize Russell's empiricist program.[11] Its origins are to be found somewhere in the works of David Hume, and at least one serious endeavor of completing this project in the modern era can be found in Ludwig Wittgenstein's *Tractatus*.[12] According to Quine, the goal at which Carnap aims is unachievable in principle, mainly because Carnap has not provided any means of translating statements of the form 'quality q is at point-instant x, y, z, t' into the statements about elementary sense experience. Hence, the crucial problem lies at the very heart of the constitutional system presented in *Aufbau*.

In a similar fashion – i.e. as a phenomenalist, reductionist project – the book has been interpreted by Goodman:

> In *Der logische Aufbau der Welt*, Carnap's purpose is to sketch a particularistic system[13] much more coherently and comprehensively than any proposed before. Starting with the few positive findings of earlier particularistic thought, he avails himself of the methods and the model of symbolic logic in developing his constructions. If the resulting system remains rudimentary and inadequate in many respects, the advance it makes must still not be overlooked.[14]

---

[10] Idem., *Ontological Relativity and Other Essays*, 73-74.
[11] Carnap even used Russell's motto at the beginning of his book: "The supreme maxim in scientific philosophizing is this: wherever possible, logical constructions are to be substituted for inferred entities."
[12] Cf. Friedman and Creath (Eds.), *The Cambridge Companion to Carnap*, 130-131.
[13] Particularistic in this context simply means a kind of phenomenalistic system, i.e. a system with phenomenal basis [RW, PL].
[14] Goodman, *The Structure of Appearance*, 151.

Goodman is honest when he admits the advancement achieved by Carnap, but, simultaneously, he stresses the elementary character of the construction in *Aufbau*. Goodman, in his critique of the book, has focused mainly on the method used there – so called quasi-analysis, and on the ontological assumptions, which, according to him, are indispensable for Carnap's project.[15] However, Goodman was not completely skeptical about the constitutional system developed in *Aufbau*. This follows from the fact that he considered his critique as a means of improving it: "[t]he purpose of my critical scrutiny is not to disparage his accomplishment, but to determine just where the remaining problems lie and perhaps to pave the way for their solution".[16]

The third of the above mentioned authors – Ayer – stressed yet another important tenet of Carnap's work. According to Ayer, but also to many other philosophers, the project undertaken by Carnap is highly anti-metaphysical in nature.[17] The verification principle states explicitly that statements which cannot be translated into statements about experiences coming from the ordinary senses are meaningless, therefore the whole metaphysical discourse, which is not expressible in phenomenalistic language remains meaningless. As Ayer puts it: "If a putative proposition fails to satisfy this principle [of verifiability], and is not a tautology, then I hold that it is metaphysical, and that, being metaphysical, it is neither true nor false but literally senseless."[18] Ayer's personal attitude towards the verifiability principle is unquestionable, but he also explains to whom he is most indebted for the promise of bringing this principle into life: "The philosophers with whom I am in the closest agreement are those who compose the 'Viennese Circle',

---

[15] Cf. Paweł Kawalec "Wstęp," [Introduction] in *Logiczna struktura świata by Rudolf Carnap*, translated by Paweł Kawalec, (Warszawa: Wydawnictwo Naukowe PWN 2011), xciv-xcvi.

[16] Ibid. However, Hannes Leitgeb noted that problems raised by Goodman are in principle indispensable, see Hannes Leitgeb: "A New Analysis of Quasianalysis," *Journal of Philosophical Logic* 36, (2007): 181-226.

[17] Quine expresses a similar view in his essay "Epistemology naturalized", where he writes: "Carnap and the other logical positivists of the Vienna Circle had already pressed the term >>metaphysics<< into pejorative use, as connoting meaninglessness" (cf. Quine, *Ontological Relativity and Other Essays*, 82).

[18] Ayer, *Language, Truth and Logic*, 41.

under the leadership of Moritz Schlick, and are commonly known as logical positivists. And, of these, I owe most to Rudolf Carnap."[19] Ayer names Carnap as his greatest inspiration, and it is rather evident that what he must have meant by this, is – among other things – Carnap's attempt to put the verifiability principle into reality.

With the above discussion, we wished to establish the facts about the long-lasting interpretation of the project that was undertaken by Carnap in *Aufbau*. All quoted authors agree with regard to the main ideas in this work. The logical positivist reading of *Aufbau* was indirectly supported by the fact that a lot of time passed before the English translation of the book was published, which allowed for a direct analysis of Carnap's work. Up to that time, works of Ayer, Quine and Goodman served as the main source of information about Carnap's project and perforce were responsible for the received view of *Aufbau*. We will now move on and describe contemporary interpretations of the book under discussion, by which we hope to show a significant change that took place at the beginning of the 1990's.[20]

4. Contemporary interpretations of *Aufbau*

The received view of *Aufbau* has become less significant recently, and this is mostly due to works of Michael Friedman and Alan Richardson. They suggested that Carnap's book should be placed in the context of Kantian or neo-Kantian thought, rather than within the empiricist tradition. Friedman puts it in the following way in the final part of one of his articles:

> The preceding analysis results in a very different picture of the epistemological significance of the *Aufbau* than that which has been customary within contemporary Anglo-American philosophy. The *Aufbau* is not best understood as starting from fundamentally empiricist philosophical motivations and then attempting to put these into effect – on the basis of the new mathematical logic of *Principia Mathematica* – in a more precise and rigorous way than had been previously possible. The epistemological motivations of the *Aufbau* begin rather

[19] Ibid., 42.
[20] Historically speaking it was Susan Haack who, in her 1977 paper, first noted the Kantian features of *Aufbau*. This remark was made by Richardson, "Logical Idealism and Carnap's Construction of the World," 60.

> with the concerns and problems of the neo-Kantian tradition: with a concern for depicting how the cognitive process transforms inherently private and subjective sensations into fully objective experience capable of validity and truth, and with the corresponding problem of carrying out this project is an essentially "logical" - that is, non-metaphysical and non-psychological – fashion.[21]

Similar conclusions can be found in Richardson's well-known book, *Carnap's Construction of the World. The Aufbau and the Emergence of Logical Empiricism*:

> [...] for Carnap the problem is how we achieve objective knowledge in the sciences, despite the subjective origin of empirical knowledge in private sensation. The key to the answer is the idea that all streams of experience have a common form. This understanding of the problem of epistemology and its solution clearly has greater affinity to the Kantian critical philosophy than it does to the foundationalism of either the traditional empiricist or rationalist camp.[22]

The quotations above clearly indicate important constituents of contemporary interpretations of *Aufbau*. It should be transparent that the discussed book was not meant as an advancement of the empiricist and reductionist program. The significance of a particular version of a constitutional system presented in the book does not stem from the fact that it is a phenomenalistic system (with an auto-psychological basis), but rather that it serves as an example of any constitutional system. To put it in other words, according to current interpretations, constitutional theory is of greater importance than any specific constitutional system. One could ask, why this is the case. The answer would be: by developing a constitutional theory, which allows the construction of any constitutional system, one can attempt to demonstrate the objectivity of all scientific statements. That attempt – very Kantian in its nature – is recognized as the central aim of the project behind *Aufbau*.

Friedman and Richardson differ slightly when they try to describe the way in which Carnap approached this central aim. According to Friedman, the objectivity of scientific statements is achieved thanks to purely structural, definite descriptions. Although our sense experiences constitute a subjective starting point, the structure of experiences remains the same for different individuals. Therefore, objectivity is

[21] Friedman, *Reconsidering Logical Positivism*, 141.
[22] Richardson, *Carnap's Construction of the World,* 28.

guaranteed by similarity of structures and the possibility of purely structural reference to every possible object investigated by science. Richardson has a somewhat different view on this matter: he claims that there are actually two projects in *Aufbau*. The first one is as just described, and the second one aims at achieving the objectivity of scientific statements by creating an intersubjective world. The constitution of this intersubjective world is possible thanks to the similarity of psychological objects of different subjects. Hence, the world of objective scientific statements is, in a sense, compiled out of many private 'frameworks' created by individuals.

The discrepancy between Friedman and Richardson results in the fact that they give different reasons why Carnap has not achieved his aim. According to Friedman, it is because purely structural definite descriptions are used to define all but one concept, but the crucial one. The central concept of the recollection of similarity relation remains undefined in the required manner. Therefore, Friedman concludes that Carnap's attempt is a failure. Richardson, on the other hand, pointed out a certain tension in *Aufbau*, caused by the presence of two projects, which are not completely compatible with each other.

5. Discussion between two views on *Aufbau*

It has been pointed out that *Aufbau* was traditionally interpreted as a realization of Russell's empiricist project. However, this view has been recently challenged, as we could observe in the previous part of our paper. Below, we present some arguments to support our claim that there has been a certain shift, and now different elements of Carnap's philosophy are being highlighted. We wish to stress this change that took place around two decades ago, because it has brought new light to research on Carnap.

The classical interpretation of *Aufbau* is justified at least to some extent. Even Carnap himself admits that in the preface to the second edition of the book:

> In this book I was concerned with the indicated thesis, namely that it is in principle possible to reduce all concepts to the immediately given [...] I wanted to attempt, for the first time, the actual formulation of a conceptual system of

> the indicated sort; that is to say, I was going to choose, to begin with, some simple basic concepts, for instance sensory qualities and relations, which are present in the raw material of experience; then I was going to formulate on this basis further definitions for concepts of various kinds.[23]

However, it has to be underlined that both Quine and Goodman read *Aufbau* in the light of the particular constitutional system that was developed by Carnap, and as we saw in the second part of our paper, the particular constitutional system presented in *Aufbau* was not Carnap's primary aim. Rather, his task was more general and consisted of creating a constitutional theory, a methodology of creating all sorts of systems, equally acceptable. The interpretation put forward by Quine and others overemphasizes the importance of part of Carnap's work, which displays only the exemplification of his broader idea.

Moreover, we wish to point to the fact, that logicism determines Carnap's work in a similar way. Let us take a look at the quotation where Carnap explains the importance of Russell's approach to philosophy:

> The new type of philosophy has arisen in close contact with the work of the special sciences, especially mathematics and physics. Consequently they have taken the strict and responsible orientation of the scientific investigator as their guideline for philosophical work, while the attitude of the traditional philosopher is more like that of a poet.[24]

This quotation indicates that what Carnap actually found the most interesting in the works of Russell was not his empiricist project, but rather the idea of scientific philosophy based on the firm ground of logic.[25]

Another issue concerns verificationism – which is often recalled when discussing *Aufbau*. The idea definitely has its origin in the Vienna Circle, and as Carnap points out, its actual development was due to Wittgenstein:

---

[23] Rudolf Carnap, *The Logical Structure of the World and Pseudoproblems in Philosophy*, translated by Rolf A. George (Chicago: Open Court 1969), vi.

[24] Ibid., xvi

[25] The Russellian inspirations of Carnap are discussed to a great extent in: Christopher Pincock, "Russell's influence on Carnap's *Aufbau*," *Synthese* 131, (2002): 1-37. It has to be noted that Carnap attended Frege's lectures in Jena between 1910 and 1914, and these were another contribution to Carnap's respect for logic.

> The most decisive development in my view of metaphysics occurred later, in the Vienna period, chiefly under the influence of Wittgenstein. I came to hold the view that many theses of traditional metaphysics are not only useless, but even devoid of cognitive content. They are pseudo-sentences... [...] The view that these sentences and questions are non-cognitive was based on Wittgenstein's principle of verifiability. This principle says first, that the meaning of a sentence is given by the conditions of its verification and, second, that a sentence is meaningful if and only if it is in principle verifiable, that is, if there are possible, not necessarily actual, circumstances which, if they did occur, would definitely establish the truth of the sentence.[26]

The principle under discussion was introduced by Wittgenstein and it was defended by him until his turn towards ordinary language philosophy. Carnap admits in this passage that he was committed to the principle of verification, but it is a historical fact that he was not the first one who came up with the idea. Therefore Carnap should be seen as the one who only makes use of the idea that was put forward and predominantly defended by others. It is of vital importance to keep that fact in mind because Carnap's anti-metaphysical attitude was absent in his writings until his *Scheinprobleme der Philosophie*[27], but at that time *Aufbau* had already been completed.

Associating Carnap's work with verificationism had major consequences. The main one was an anti-metaphysical reading of the theses promoted by him in *Aufbau*. It is clearly present in Ayer, as we have already seen.[28] Carnap's views on metaphysics at the time of publication are not entirely transparent. However, when one considers his general aim (as discussed above) and the appropriate sections of *Aufbau* (§§ 12, 16), then it is at least possible to accept his neutrality with respect to metaphysics. The language of the constitutional system is neutral, and serves equally well for realistic and idealistic formulations of theses. According to Carnap, the language of logic is universally applicable and has great expressive power, which makes it a perfect means for scientific philosophy.

---

[26] Rudolf Carnap, "Intellectual Autobiography," in *The Philosophy of Rudolf Carnap*, edited by Paul Arthur Schilpp (Chicago: Open Court, 1963), 45.
[27] Published in 1928.
[28] Ayer, *Language, Truth and Logic*, 41.

Furthermore, Carnap's general aim is revealed by the analysis of the story behind the choice of the title for his book.[29] First of all, it has to be noted that Carnap planned to write two books: one presenting the phenomenalistic constitutional system - this is the one he actually wrote - and another one which would contain the physicalistic constitutional system. In both cases, the systems being developed served only as an exemplification, rather than as an expression, of the actual beliefs of Carnap with regard to the constitution of the world. There is a high dose of contingency in the fact that Carnap first wrote a book containing the phenomenalistic system, and that he never completed his second book. What is more, it was actually the second book that was supposed to be titled *Der logische Aufbau der Welt*, while the first one was supposed to be titled *Der logische Aufbau der Erkenntnis*. This indicates that the title refers only to the particular constitutional system presented in the book.[30]

The final point that we want to raise stems from the work by Friedman and Richardson. Their thorough investigations show that there is strong Kantian import in Carnap's early work. The main question, to which he was seeking an answer, was: how do we obtain objective scientific statements about the world, considering a subjective starting point, i.e. our sense experiences? According to Friedman and Richardson (although their interpretations are somewhat different), only if we focus our investigations around the above formulated question, will we be able to fully appreciate and adequately comprehend the ideas hidden behind *Aufbau*.

We do not wish to say that one interpretation of *Aufbau* is better than the other. Our aim is not to go so far. We limit ourselves only to broadening the perspective. Both classical and modern interpretations of *Aufbau* constitute the extremes of a broader spectrum of approaches to

---

[29] Cf. Alberto J. Coffa, *The Semantic Tradition from Kant to Carnap: The Vienna Station*, (New York: Cambridge University Press 1991); Richardson, "How not to Russell Carnap's Aufbau," 10.

[30] An even more direct indication comes from private correspondence between Carnap and Schlick from 1927. The title that Carnap was considering at that time was: *Der logische Aufbau der Welt. Versuch einer Konstitutionstheorie der Begriffe* [The logical structure of the world. An attempt at the theory of constitution of concepts], cf. Coffa, *The Semantic Tradition from Kant to Carnap: The Vienna Station*, 231.

the book. Carnap found inspiration for his ideas in a variety of different sources, ranging from Gottlob Frege and Russell's logicism, to Ernst Cassirer and Heinrich Rickert's neo-Kantianism. Therefore, the scope of possible interpretations of *Aufbau* is equivalently broad.

Literature

Ayer, Alfred J. *Language, Truth and Logic*. New York: Penguin Books, 1978.

Carnap, Rudolf. "Intellectual Autobiography." *In the Philosophy of Rudolf Carnap*, edited by Paul Arthur Schilpp, 3-86. Chicago: Open Court, 1963.

——— *The Logical Structure of the World and Pseudoproblems in Philosophy*. Translated by Rolf A. George. Chicago: Open Court, 1969.

Coffa, Alberto J. *The Semantic Tradition from Kant to Carnap: The Vienna Station*. New York: Cambridge University Press, 1991.

Friedman, Michael. *Reconsidering Logical Positivism*. New York: Cambridge University Press, 1999.

——— "The Aufbau and the Rejection of Metaphysics." In *The Cambridge Companion to Carnap*, edited by Michael Friedman and Creath Richard, 129-152. New York: Cambridge University Press, 2008.

Friedman, Michael and Creath, Richard (eds.). *The Cambridge Companion to Carnap*. New York: Cambridge University Press, 2008.

Goodman, Nelson. *The Structure of Appearance*. Cambridge: Harvard University Press, 1977.

Haack, Susan. "Carnap's Aufbau: Some Kantian Reflexions." *Ratio 19* (1977): 170-175.

Kawalec, Paweł. "Wstęp" *[Introduction] in Logiczna struktura świata by Rudolf Carnap*. Translated by Paweł Kawalec, xi-cviii. Warszawa: Wydawnictwo Naukowe PWN, 2011.

Leitgeb, Hannes. "A New Analysis of Quasianalysis." *Journal of Philosophical Logic 36* (2007): 181-226.

Pincock, Christopher. "Russell's influence on Carnap's Aufbau." *Synthese 131* (2002): 1-37.

Quine, Willard V. O. *From a Logical Point of View*. New York: Harper Torchbooks, 1963.

——— *Ontological Relativity and Other Essays*. New York: Columbia University Press, 1969.

Richardson, Alan. "*How not to Russell Carnap's Aufbau.*" In PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1990 (1990): 3-14.

——— "Logical Idealism and Carnap's Construction of the World." *Synthese 93* (1992): 59-92.

——— *Carnap's Construction of the World. The Aufbau and the Emergence of Logical Empiricism*. New York: Cambridge University Press, 1998.

Schilpp, Paul Arthur (ed.). *The Philosophy of Rudolf Carnap*. Chicago: Open Court, 1963.

Stroud's Dissatisfaction with Kant: On the Indispensability of "Certain Ways of Thinking"

Elvira Basevich
(The Graduate Center, CUNY)

In his recent book, *Engagement and Metaphysical Dissatisfaction*, Barry Stroud criticizes Kant's contention that certain indispensable ways of thinking, which are necessary features of the a priori structure of human cognition and experience, generate insights into the objective nature of the external world. Because the a priori features of human cognition necessarily correspond to the external world, Stroud argues that Kant achieves an undue metaphysical "satisfaction." As part of his transcendental idealism, Kant simply stipulates that the a priori structure of human understanding and experience corresponds to the objective structure of the external world. This essay argues that Stroud mischaracterizes Kant's epistemological project in the *Critique of Pure Reason*. Part of Kant's critical grounding of theoretical reason involves demonstrating that the a priori structure of human cognition and experience does not reveal any truths about the fundamental nature of objective reality; only sensible intuition furnishes us with a limited insight into objective reality. This essay shows that Stroud a) attributes an unjustifiable metaphysical realism to Kant that contends that one can "read off" the objective structure of the external world from the a priori structure of human cognition and experience, and b) ignores the role of sensible intuitions in Kant's account of knowledge.

1. Introduction

Barry Stroud's recent book, *Engagement and Metaphysical Dissatisfaction*, analyses the aspirations of metaphysics in the history of Western philosophy. Stroud argues that philosophers cannot ground metaphysics on a stable foundation because it is impossible to justify a metaphysical position – any metaphysical position – with any certainty. For this reason, metaphysical inquiry is bound to lead a philosopher to "dissatisfaction." Stroud conceives metaphysics as the study of the fundamental nature of reality and emphasizes that it is primarily an epistemological project. He adds the important proviso that reality as it exists independently of the subject is the object of metaphysical inquiry.

This poses the question of how – and whether – a subject with a limited and partial perspective of the external world can have objective knowledge of it. For, there is a distinction between knowledge of reality and reality as it is in itself. Broadly speaking, then, his question is: What epistemic warrant do we have to affirm *any* metaphysical truth, if we lack a sufficiently impartial epistemic point of view simply by virtue of being finite sensible subjects?

Whether a finite sensible subject can have access to the fundamental nature of reality, and thus acquire justified beliefs, is a perennial question of Western philosophy. It is also Kant's question in his *Critique of Pure Reason*. There, Kant attempts to demonstrate that knowledge is not only possible, but also firmly grounded. He, however, abandons the hope that finite sensible subjects can have access to reality as it is in itself, understood as some metaphysically real entity existing completely independently of the subject. As a consequence, metaphysical truths are based on appearances relative-to-us rather than on the thing-in-itself. The a priori structure of the human mind generates the necessary conditions for human knowledge and provides the cognitive framework by virtue of which phenomenal objects appear to us. That the conditions for knowledge are dependent on the a priori structure of the human mind does not count against the "objectivity" of metaphysical knowledge. Rather, those conditions serve to *constitute* the epistemic criterion of objectivity. Part of that criterion dictates that critical theoretical reason must ultimately rely on sensible intuitions for its content. The thing in-itself, which is beyond our sensibility, is relegated outside the scope of justifiable metaphysical knowledge. Although objectivity is an epistemic criterion dependent on the a priori structure of the human mind, it does not render metaphysics hopelessly partial. The charge of partiality only makes sense in the context of a feasible alternative, one that shows how a finite sensible subject can grasp the thing-in-itself. Kant claims that such an alternative is epistemically incoherent.

Stroud is not impressed with Kant's method for finding metaphysical truth, or, as he puts it, "satisfaction." He believes that Kant wrongly contends that the necessary conditions for human knowledge and experience shed light on the fundamental nature of reality. He argues that Kant achieves an unjustified metaphysical satisfaction by

asserting that the principles, beliefs, and concepts that are indispensable for human thought necessarily represent objective reality. For Stroud, the indispensability of "certain ways of thinking" is precisely the obstacle for grounding metaphysical inquiry on a stable epistemic foundation.[1] He defends his position with the philosophical truism that the fact that a particular human being thinks something – or even that all human beings think something – does not reveal any truth about objective reality.

In criticizing Kant, Stroud's principal mistake is supposing that for Kant the mere indispensability of "certain ways of thinking" (e.g., principles, beliefs, concepts) reveals insight into the fundamental nature of objective reality. Stroud is imprecise in his use of the terms "belief," "principles" and "propositions," which he groups together as "certain ways of thinking." Kant's metaphysical project in the *Critique of Pure Reason* distinguishes the a priori forms of sensibility from the a priori categories. These two elements serve very different functions in Kant's epistemology and, taken together, do not reveal any necessary truths about objective reality as it is in itself. This essay argues that, first, part of Kant's critical grounding of theoretical reason involves demonstrating that theoretical reason does *not* reveal any truths about objective reality as it is in itself; only sensible intuition furnishes us with limited insight into it. This is part of Kant's metaphysical modesty.[2] In this sense, Stroud shares the *same* concern with Kant about the uncritical use of theoretical reason in making judgments about objective reality. Second, Stroud ignores the role of sensible intuitions in human knowledge, whereas Kant takes great pains to illustrate the importance of external sensible inputs for knowledge in the *Critique of Pure Reason*.

In the first section, I outline Stroud's criticism of Kant's epistemological project and focus on his contention that Kant relies on "certain indispensable ways of thinking" in attempting to gain insight into objective reality.[3] In the second section, I present Kant's distinction

[1] Barry Stroud, *Engagement and Metaphysical Dissatisfaction: Modality and Value* (Oxford: Oxford University Press, 2011), 125.
[2] This brings his criticism of Leibniz to the fore.
[3] I use the term 'objective reality' the same way as Stroud does. For him, objective reality means a totally subject-independent external world --- equivalent to the Kantian thing-in-itself.

between the epistemic roles of the understanding and sensibility and show how Stroud's interpretation misses Kant's metaphysical modesty. In the concluding section, I show some of the epistemological and metaphysical implications of Kant's project in the *Critique of Pure Reason* in light of Stroud's criticisms. In illuminating how Stroud mischaracterizes Kant's project in the *Critique of Pure Reason*, I hope to demonstrate that that latter is more plausible and coherent than is often supposed in contemporary analytic Anglo-American philosophy. At the very least, I hope to show what sorts of philosophical assumptions one cannot make when criticizing Kant's metaphysical and epistemological theses in the *Critique of Pure Reason*.

2. Stroud's Criticism of Kant on the Indispensability of 'Certain Ways of Thinking'

Stroud argues that for Kant the indispensability of certain ways of thinking is "the key to the very possibility of putting metaphysics once and for all on the secure path of a science."[4] He rejects this starting point for metaphysics by claiming that the indispensability of certain ways of thinking is more likely to lead one to uncertain, rather than certain, knowledge.[5] According to Stroud, Kant holds that human thought and experience reveals insights into objective reality and thus provides us with indubitable metaphysical truths. Since the necessary conditions of thought and experience must correspond to a subject-independent, objective reality, one can "deduce" from these conditions truths about objective reality.[6] That is, given that certain cognitive elements must obtain in human cognition in general, they too must obtain when on thinks about the subject-independent external world. Stroud attacks Kant by asking a straightforward question: What is the relationship between our indispensable ways of thinking and objective reality?

Stroud separates the question of what sort of things necessarily constitute human thought and experience from the metaphysical

[4] Stroud, *Engagement*, 126.
[5] Stroud, *Engagement*, 143.
[6] Stroud, *Engagement*, 130.

question concerning the nature of objective reality.[7] According to him, Kant's mistake is to assume that there is an essential connection between the former and the latter. He interprets Kant as claiming that the elements necessary for human cognition disclose objective reality *because* these elements are indispensable for human cognition. Given that human cognition must take a certain form, our thoughts about and experiences of the external world must conform to the general pattern necessary for human cognition. A question nonetheless remains for Stroud about what the world is "independently of us and our responses to it,"[8] one he thinks Kant does not take sufficiently seriously.

Stroud's analysis of the relationship between human cognition and objective reality within the Kantian framework assumes that the mere existence of thought and experience discloses something about objective reality, including questions about causality, substance, space and time, and other "indispensable" cognitive elements. He claims that Kant's deduction, if successful,

> would show...that there are enduring objects in space and time, that all events stand in causal relations to other events that happen, that there are thinking, experiencing, active subjects who hold these and other such general beliefs about the world. If Kant is right, all these things must be so if anyone thinks or has experiences of anything at all; they are among the necessary conditions of all thought and experience. They state the way things must be, given that there is any thought and experience.[9]

So, Stroud contends that certain "things...must be so" if anyone has any thoughts or experiences. There is an obvious ambiguity in his phrasing of the relationship between what he takes to be transcendentally real objects – "things" – and the necessary conditions for thought and experience. Stroud is especially unclear in his characterization of objective reality, referring to it simply as "the way things must be." He seems to think that Kant's epistemology, which analyzes the a priori conditions necessary for thought and experience, also commits Kant to the independent metaphysical status of those conditions in the external

[7] As I will show, separating these two is a mistake, although it is crucial for Stroud's argument.
[8] Stroud, *Engagement*, 125.
[9] Stroud, *Engagement*, 130.

world.[10] He interprets Kant as saying that we have access to these metaphysical truths "if anyone thinks or has experiences of anything at all" and, at the same time, those truths hold entirely independently of the subject, so long as there is someone thinking something. "The necessary conditions of all thought and experience" somehow guarantee the independent metaphysical status of cognitive elements that are indispensable for human thought and experience. The implication of this reading of Kant is that the indispensability of certain beliefs entails a set of robust metaphysical commitments. Stroud apparently supposes that Kant's epistemology is a peculiar type of correspondence theory that guarantees the independent "metaphysical" status of the elements indispensable for our cognition. For him, "the full resources of Kantian metaphysics...would reach beyond indispensability to the relation between the contents of those indispensable beliefs and what is so in the world."[11] But how does Kant guarantee that certain indispensable ways of thinking necessarily correspond to the external world?

Stroud thinks Kant's solution to this problem was to introduce idealism. In order to defend the claim that our indispensable beliefs hold in the external world, Kant has to defend idealism as a metaphysical doctrine. Idealism allows for an essential and necessary correspondence between our beliefs and the external world. "To conclude that the beliefs must represent the way things *really are* because they are beliefs that any thinker must accept sounds like the familiar but hollow arrogance of idealism: things must be a certain way because we thinkers must think things are that way."[12] Stroud does not see that this reading of Kant turns Kant into a Berkeleyan subjective idealist, for whom reality is mere thought, and thus generated by a grand mind. For Stroud, the only way to guarantee the necessary correspondence between human cognition and reality is to assume that the two are essentially identical. Our cognition corresponds to the external world because the external world is just made up of thought.

The foundation of the external world for Kant is not an omniscient and omnipresent mind of god à la Berkeley; this is not how knowledge

[10] Stroud, *Engagement*, 132.
[11] Stroud, *Engagement*, 139.
[12] Stroud, *Engagement,* 143. My emphasis.

of the external world is grounded for him. As I will show in the next section of this essay, such a position is untenable for Kant because knowledge would have to rely on the intellectual intuition of an omniscient mind, the truth of whose judgment is not reliant on sensory inputs. The genius of Kant is that he does *not* comment on the ultimate metaphysical status of objects, let alone provide their foundation. In transcendental idealism the foundation of epistemology is grounded in the a priori structure of the human mind, not in the transcendental reality of (god's) thought. In approaching metaphysics through a delineation of our transcendental psychology, Kant hopes to avoid having to make any sort of commitments about the necessary features of the external world. He can thus limit himself to describing the necessary features of human cognition and experience without stipulating that these features are in any way necessary attributes of the external world.

Moreover, Stroud fails to realize that his interpretation of Kant's idealism excludes *truly* external sensory inputs. He seems to think that the content of the categories and, more importantly, of our sensibility is already implicit in the mind. Anything we can assert about the external world is already contained in the a priori features of our cognition. The mere existence of human thought and experience "yields substantial[13] conclusions about the world."[14] This follows naturally given that Stroud thinks that for Kant the external world is mere thought. Quite understandably, having minimized the role of sensible intuitions and unwittingly ascribed a transcendental realism (in the form of a Berkleyan omnipresent and omniscient mind) to Kant, Stroud questions whether the conditions indispensable for human thought and experience establish the fundamental nature of the external world. He asks:

> Can we proceed by secure necessary steps from our thinking and experiencing things in certain ways, or even from our thinking and experiencing anything at all, to conclusions about how things actually are whether they are thought about or experienced or not? Can you get a whole independent world out of nothing

[13] Stroud also uses the word "substantive" in other parts of his book to describe the kind of conclusions he has in mind here.
[14] Stroud, *Engagement*, 131.

> more than thought of a world or a world-like course of experience? Can you get a rabbit out of a silk hat?[15]

What Stroud fails to realize is that Kant's answer, like his own, to all these questions is a resounding 'No!' Like Stroud, Kant holds that "the world carries no implications about thinkers thinking anything."

3. How to Get Kant Wrong: Conflating Understanding and Sensibility

Emphasizing the indispensability of certain elements of human cognition in general is the wrong approach for understanding Kant's project in the first *Critique*. Kant does not think that certain ways of thinking generate metaphysical truths about the fundamental nature of external reality, unless one is a perfect being privy to intellectual intuitions. Stroud attributes a metaphysical realism to Kant that the latter rejects. In this section, I will show how Stroud's interpretative approach fails to distinguish the distinct epistemic roles of the pure forms of sensibility and the categories of the understanding. Such an approach ignores Kant's overarching project of checking the uncritical use of theoretical reason in making metaphysical claims. A central part of Kant's project is to show that human cognition does not proffer insight into objective reality as it is in itself. In the following section, I delineate some implications of Kant's epistemological and metaphysical project in context of Stroud's criticism.

To be sure, Stroud is right to emphasize the role of human cognition in Kant's metaphysics. As Allen Wood writes, "Kant understood the term metaphysics in epistemological terms. That is, for the purposes of metaphysics, 'nature' is what is known through experience, and so 'metaphysics' is a science demarcated not by a set of objects with which it deals but by the a priori epistemic status of its principles."[16] Yet, the perception of objects that are given in experience is nonetheless a crucial feature of Kant's epistemology. Kant thinks that

[15] Stroud, *Engagement*, 132.
[16] Allen M. Wood, *Kant: Blackwell Great Minds* (Malden, Oxford, and Carlton: Blackwell Publishing, 2005.), 24.

his emphasis on perception dovetails with his conception of the formal, a priori framework of metaphysical knowledge. In failing to distinguish the role of the understanding from sensibility, Stroud minimizes the contribution of sensible intuitions to knowledge.

Kant begins the *Critique of Pure Reason* by underscoring the "fact of experience" (*Faktum der Erfahrung*). He argues that two main faculties are critical for knowledge: the passive sensibility and the active understanding. Both are parts of the a priori structure of the human mind and could be isolated independently of experience; in fact, our sensibility precedes any particular perception and ultimately makes experience itself possible for us.[17] But this does not mean that the mind alone somehow "generates" perceptions imminently, as Stroud's interpretation suggests. On his reading of Kant's idealism, the mind – presumably of god – gives human beings their sensible intuitions.

In the Transcendental Aesthetic, however, Kant defines sensibility as "the capacity (receptivity) to obtain representations through the way in which we are affected by objects in the senses."[18] The capacity to be affected by objects via perception is part of the a priori structure of our minds. Kant calls this the pure form of sensibility. Insofar as we are capable of passively receiving sensory stimuli and having experiences, our mind has "immediate" and "direct" reference to external objects;[19] this is "at which all thought aims."[20] Although our pure form of sensibility makes it possible for us to receive sensory inputs, it does not "constitute" those inputs; those inputs are given from without. "Objects are therefore given to us *by means of* our sensibility."[21] Although our sensibility delimits the form our experiences can take, our sensibility does not "determine" the content of sensory inputs in perception nor does it entail any robust metaphysically commitments. For, as we shall see, having an experience of an appearance in space and time does not entail affirming the transcendental realism of space and time. Nor are we warranted in attributing a transcendental realism to the objects to

[17] Kant, Immanuel, *Critique of Pure Reason*, translated and edited by Marcus Weigelt (London: Penguin Group, 2007), A24/B39.
[18] Kant, *Critique of Pure Reason*, A19/B33.
[19] Kant, *Critique of Pure Reason*, A21/B35.
[20] Kant, *Critique of Pure Reason*, A19/B33.
[21] Kant, *Critique of Pure Reason*, A19/B33. Emphasis added.

which we have immediate and direct access in perception.

Sensory inputs, which are given to our sensibility, are the indispensable starting points for metaphysical knowledge. Kant claims that without them knowledge is empty.[22] He is adamant about this point, arguing that "things, which cannot be given to us by any experience, are nothing for us."[23] The understanding is that by means of which we "think" a sensible intuition, but they "by themselves" cannot "yield knowledge of objects."[24] On the other hand, a sensible intuition without the categories would not be cognizable by our minds. Sensible intuitions occasion the application of the pure principles of the understanding.

In the absence of sensible intuitions, "the understanding runs the risk of making, through mere sophism, a material use of the merely formal principles of the pure understanding, and thus making judgments indiscriminately about objects which are not given to us, nay, which perhaps can never in any way be given."[25] The "material use" of the "formal principles of the pure understanding" is the mistake he attributes to Leibniz and his followers. They tried to derive substantial conclusions about the world by appealing to nothing other than logical consistency and the principle of non-contradiction; that is, they appealed to nothing other than thought. With the exception of moral reasoning, the "empirical use of the understanding" is thus the only appropriate use of it.[26]

Moreover, Kant does not think that the general form of human cognition entails any robust metaphysical commitments. One cannot say in any "satisfying" sense that space and time are transcendentally real or that the principle of causality objectively obtains. With regard to the latter, for example, what one can say is that in order for experience to have a coherent form, where one event follows another with necessity, the understanding applies the pure principle of causality to appearances. Necessity is never empirically perceived or manifest as some sort of metaphysical property of empirical objects. Part of Kant's metaphysical

[22] Kant, *Critique of Pure Reason*, A63/B87.
[23] Kant, *Critique of Pure Reason*, A50/B74.
[24] Kant, *Critique of Pure Reason*, A19/B33.
[25] Kant, *Critique of Pure Reason*, A63/B88.
[26] The exception here is moral considerations.

modesty is his refusal to ascribe objective reality to the principle of causality that extends beyond the phenomenal world of appearances. Even in its application to the phenomenal world, the principle of causality (and substance) conceptualizes given sensory inputs without positing the principles in a transcendentally real sense. But this does not mean that external reality is a systematically organized illusion of the human mind, since immediate sensory inputs are required for the instantiation of the principle of causality or any of the categories in the first place.[27] Phenomenal appearances are thus perfectly congruent with empirical realism. "Experience is not limited to the private domain of our own representations, but includes an encounter with 'empirically real' spatiotemporal objects."[28] In fact, experience limited to the private domain would be impossible, for it must begin with a sensible input that is given from without. Appearances are at the same time metaphysically "hollow" in that they lack a fundamental metaphysical foundation outside the formal, a priori structure of our sensibility and understanding.

4. Some Implications of Kant's Epistemological and Metaphysical Project

Now that we have a sense of the role of sensibility in Kant's epistemology, we can begin to see that Kant's idealism is not nearly as "arrogant" as Stroud seems to think. Objects do not necessarily exist because of certain indispensable aspects of human thought and experience. Stroud's account of transcendental idealism countenances that metaphysically real "things" essentially and necessarily correspond to certain indispensable elements of human cognition. Necessity, however, is a mark of the a priori and thus reflects the structure of the mind. That empirical objects necessarily appear to us in a certain form and under certain conditions does not mean that those objects themselves possess any necessary features or even exist as subject-

[27] Kant, *Critique of Pure Reason*, B69.
[28] Henry Allison, "Kant's Transcendental Idealism: Chapters 1 and 2," in *Kant's Critique of Pure Reason: Critical Essays*, edited by Patricia Kitcher, 181-217. (Lanham, Boulder, New York and Oxford: Rowman and Littlefield Publishers, Inc., 1998), 185.

independent transcendentally real "things" bearing those features. This is exactly the kind of metaphysical commitment Kant wants to avoid making. For "in relation to the merely subjective conditions [of our sensibility] all objects are appearances and do not exist *for themselves* as things given to us in this way."[29] Contextualizing objects in relation to our a priori epistemic framework, which includes the understanding and the form of our sensibility, prevents us from ascribing any properties to things-in-themselves.

Sensible intuitions "are *not* such that through it itself the existence of the object is given (such an intuition, so far as we can understand, can belong only to the primordial being)."[30] The empirical use of the understanding does not disclose anything about reality in itself. Even the transcendentally real "existence of the object" as *we* conceptualize it is not taken for granted. This means that, contra Stroud, the necessary conditions for thought and experience do not correspond to transcendentally real "things." For Kant, we do not, as Stroud suggests, "proceed by secure necessary steps from our thinking and experiencing things in certain ways, or even from our thinking and experiencing anything at all, to conclusions about how things actually are whether they are thought about or experienced or not." Kant never contends that we have the epistemic license that would allow us to make these kinds of claims nor does his transcendental idealism wed him to such claims.

As Michelle Grier puts it, Kant does not have an ontology, the latter being exactly what Stroud attributes to Kant.[31] She points to a passage where Kant argues that ontology is replaced by the categories of the Transcendental Analytic. In contrast to previous philosophers who in delineating the categories of Being at the same time affirmed their transcendentally real status, Kant limits himself to explaining and justifying them in their universal and necessary application to sensible intuitions. This is all that is required to ground knowledge. As part of the a priori structure of the human mind, the categories are universal and necessary, but there is no sense in which the existence of objects, as,

---

[29] Kant, *Critique of Pure Reason*, B72.
[30] Kant, *Critique of Pure Reason*, B72.
[31] Michelle Grier, "Kant's Critique of Metaphysics," *The Stanford Encyclopedia of Philosophy* (Summer 2012 Edition), Edward N. Zalta (Ed.), URL = <http://plato.stanford.edu/archives/sum2012/entries/kant-metaphysics/>

for example, their appearance in space and time, is necessary. Stroud seems to think that in employing indispensable ways of thinking we have access to what is necessarily so.[32] According to Kant, finite, sensible beings cannot know this. What necessarily obtains in the phenomenal world in relation to the pure form of our sensibility and understanding is not what obtains in the noumenal world of the thing-in-itself.

Stroud might question why the pure forms of sensibility and the understanding are not arbitrary starting points for epistemic inquiry. If we want to gain knowledge of the external world, what is the epistemic value of affirming a starting point that is totally arbitrary from an "objective" point of view? For, after all, we could have acquired in the course of evolution an entirely different sensibility. Why assume that appearances in space and time are the only ways to conceive the phenomenal world? The pure form of our sensibility just happens to rely on spatial and temporal appearances necessarily, but this does not shed any light onto the fundamental nature of the external world. Nor does it preclude the possibility that knowledge could take other forms or that other types of sensibilities could receive appearances that are not in space and time. Kant says as much. He is emphatic that he is talking about "the *human* standpoint" and that it is only from this standpoint "that we can speak of space, extended objects, etc."[33] Because space and time are transcendentally ideal, that is, they reflect the pure form of the sensibility that human beings happen to have, Kant is comfortable with the idea that other beings might exist with a different form of sensibility. There might be a multiplicity of possible sensibilities, of which we happen to have one – that of space and time. This shows that Kant cannot allay a deep skepticism that questions why our sensibility must take the form of space and time or why only a certain set of categories can conceptualize sensory inputs. These questions are unanswerable for Kant.

Before we grow too disheartened with Kant, however, let's evaluate Stroud's alternative model of impartial knowledge. The charge of epistemic partiality presupposes a model of impartial knowledge that attempts to gain access to reality independent of any form of sensibility

[32] Stroud, *Engagement*, 133.
[33] Kant, *Critique of Pure Reason*, A27/B43. Emphasis added.

and even understanding. Stroud admits that perhaps aspects of our cognition are necessary, but this counts against our capacity to have insight into objective reality. With regard to their ultimate metaphysical status, he thinks that as finite sensible creatures we do not have enough information to judge one way or the other: perhaps our "indispensable beliefs" are features of objective reality, perhaps they are not. This is why metaphysics is bound to end in dissatisfaction. He prefers to avoid passing a positive or a negative metaphysical verdict and, in effect, relegates metaphysics to an amusing pastime we cannot seem to do without, but whose projects are doomed from the start.

Stroud can avoid this cynical outcome if he reevaluates his model of epistemic impartiality. Kant avoids this metaphysical aporia by rejecting from the start certain assumptions that Stroud makes. Stroud tacitly assumes that one ought to be able to evaluate "objectively" the indispensable elements of human cognition and by stepping outside one's self and comparing them to things-in-themselves. Affirming this ideal epistemic perspective is a key assumption of his criticism against Kant. The problem is that even according to Stroud it is impossible to define what this ideal amounts nor could we ever achieve it:

> If we could never detach ourselves even temporarily from our acceptance of indispensable beliefs..., we could not subject them to the kind of scrutiny that an impartial metaphysical assessment of their relation to reality would seem to require....Beliefs that are indispensable to any thinker's conception of a world would have to be retained as part of any conception that had been put under metaphysical assessment and survived. All indispensable ingredients of that original conception would necessarily survive any such alleged assessment.[34]

Stroud unjustifiably assumes that an ideal, godlike subject could simply compare human cognition with the subject-independent external world. Kant, unlike Stroud, is ultimately agnostic about whether such a transcendental world even exists. He contends that affirming such a standard of knowledge for finite, sensible subjects is unwarranted. Moreover, Stroud fails to articulate *on what grounds* the ideal epistemic subject would evaluate our indispensable beliefs. In what respect would their evaluation be sufficiently impartial? What object would one have to have in mind when one undertakes such an impartial metaphysical

[34] Stroud, *Engagement*, 141.

assessment? To say simply reality itself is not enough. For even a god would have to have some type of method (e.g., intellectual intuition) for achieving insight into reality. Stroud does not bother to articulate what ideal insight, independent of all forms of sensibility and human understanding, would amount to.

For Kant, on the other hand, an ideal of knowledge that is totally independent of a subject – whether sensible or divine – would completely lack content:

> Now a general [impartial] criterion of truth would be one that is valid for all knowledge, whatever its object may be. But it is clear that such a criterion abstracts from all contents of knowledge (reference to its object), while truth concerns these very contents. It is impossible and absurd, therefore, to ask for a sign of truth of such contents, and therefore a sufficient and at the same time general indicator of truth cannot possibly be given...[O]f the truth of knowledge, as far as its matter is concerned, no general indicator can be demanded.[35]

Kant's point here is that a Stroudian model of impartiality demands a "general indicator" of truth, which "abstract[s] from all contents of knowledge." It is precisely this general indicator that cannot be given. Stroud is comfortable with the fact that we will never find such a general indicator of truth and nonetheless continues to affirm it as a desirable model for metaphysical knowledge. But why hold on to such a model of truth in the first place, especially if, as Stroud argues, it leads to dissatisfaction? Stroud leaves this question unanswered.

Literature

Allison, Henry. "Kant's Transcendental Idealism: Chapters 1 and 2." *Kant's Critique of Pure Reason: Critical Essays*. edited by Patricia Kitcher. Lanham, Boulder. New York and Oxford: Rowman and Littlefield Publishers Inc., 1998.

Grier, Michelle. "Kant's Critique of Metaphysics." *The Stanford Encyclopedia of Philosophy* (Summer 2012 Edition). edited by Edward N. Zalta. URL = <http://plato.stanford.edu/archives/sum2012/entries/kant-metaphysics/>

Kant, Immanuel. *Critique of Pure Reason*. translated and edited by Marcus Weigelt. London: Penguin Group, 2007.

Stroud, Barry. *Engagement and Metaphysical Dissatisfaction: Modality and Value*. Oxford: Oxford University Press, 2011.

[35] Kant, *Critique of Pure Reason*, A60/B84.

Wood, Allen M. *Kant: Blackwell Great Minds.* Malden, Oxford, and Carlton: Blackwell Publishing, 2005.

I Can't Get No (Metaphysical Dis-)satisfaction! – Why I'll be satisfied with my metaphysical conquests

Carly Minsky
(University of Oxford)

Stroud's theory of metaphysical dissatisfaction ostensibly presents a controversial and troubling conclusion to the effect that specific metaphysical inquiries we pursue are inevitably unsatisfiable, since we are necessarily unable to justifiably endorse a positive or negative commitment about the existence of the metaphysical phenomena in question. I argue that Stroud has not sufficiently explicated the notion of inconsistency that he claims undermines an endorsement of a negative metaphysical conclusion, nor has he shown that, when this inconsistency obtains, endorsing a negative metaphysical conclusion does not satisfy the aims of the inquiry. I further argue that metaphysical dissatisfaction is not inevitable, since Stroud's conditions on metaphysical satisfaction are too strong in requiring metaphysical conclusions to be independently supported. On this basis, I reject Strouds' claim of pervasive metaphysical dissatisfaction.

1. Introduction

In successive books, Stroud (2001, 2011) claims that there are specific metaphysical inquiries that are inevitably doomed only to provide dissatisfactory answers, if any at all. These questions characterise a large part of the 'metaphysical project', namely, the part that is concerned to establish *what the world is really like* by understanding the relation between the world *itself* and our conception of the world. He supports this claim by showing that for each of the specific subject areas (e.g. causality, modality and value), the inquiry into whether these phenomena independently exist in 'objective reality' result in a dilemma: *either* we can conclude a 'negative metaphysical result' that the object in question is only a feature of our *subjective* conception of the world, but only at the cost of consistency, or else we can accept a positive metaphysical conclusion that the object in question objectively exists simply because the negative conclusion is unavailable and the inquiry was never genuinely

open. Both horns of the dilemma are *dissatisfying* with respect to our aims when we pursue these metaphysical questions. Therefore, Stroud concludes, we cannot avoid *metaphysical dissatisfaction* since the aims of a significant portion of our metaphysical inquiry are unsatisfiable from the outset.

I have some pressing questions that arise in response to Stroud's project and its conclusion. First and foremost, what exactly is the inconsistency he claims prevents us from endorsing negative metaphysical conclusions? Should we take the inconsistency to *always* undermine the satisfaction of metaphysical projects of this kind? Even if we can neither consistently endorse nor reject the *negative* metaphysical conclusions, might we nonetheless overcome *this* dissatisfaction by providing support for the *positive* metaphysical conclusion and provisionally accepting it instead? In general, how significant is Stroud's metaphysical dissatisfaction, if indeed his conclusion stands up?[1]

I will address these questions in the following paper, ultimately arguing metaphysical dissatisfaction should not be taken too seriously for three reasons: first, there are persistent question about what kind of inconsistency could and does undermine the acceptance of negative metaphysical conclusions; second, there is reason to think that in some cases, metaphysical satisfaction can be achieved by endorsing negative metaphysical results in spite of the inconsistency; thirdly, there is potential to argue *for* positive metaphysical conclusions even if we cannot rule out the negative conclusions, since our fundamental partiality towards positive metaphysical conceptions should not, on its own, undermine metaphysical satisfaction.

2. Stroud's Metametaphysics

The part of metaphysics that is the object of Stroud's *meta*metaphysical scrutiny is the project of reflecting on beliefs about the world that are taken for granted, a 'shared human conception of the world' in order to

[1] There are many other questions I can't address here. For example, is it true that positive conceptions of causality, modality and value are indispensable to human thought? And if so, is it also true that this necessarily renders negative metaphysical conclusions available only at the cost of consistency? In what follows I grant Stroud these claims for the sake of argument.

establish whether this conception '*goes beyond anything that is strictly speaking in reality*'[2]. This project aims at refining our world-view by correctly identifying which of our beliefs capture the way the world *actually* is in an independent, objective reality. Stroud assesses this metaphysical project:

> ...I want to investigate this kind of metaphysical reflections to see how its conclusions are to be reached and what support can be found for them... My main question is whether or how any active, engaged human person...can carry out a reflective philosophical project like this and arrive at metaphysical conclusions he or she can believe and find illuminating.(Stroud 2011:4-5)

In identifying the starting point of such a metaphysical project as the beliefs people *do in fact have* about the world, Stroud is intentionally advocating a loosely-Kantian metaphysical approach: '*metaphysics from within*' (*ibid*: 7) - metaphysics which investigates the conditions for us having the beliefs about the world that we want to subject to metaphysical scrutiny. However, whilst Kant prescribed this feature of metaphysics as essential to achieving metaphysical satisfaction, Stroud endorses this methodological approach as the only way metaphysics could be done, and ultimately argues it is inevitably *dis*satisfying. He also contrasts his metametaphysics with that of the logical positivists, who viewed metaphysics as unsatisfiable *by its very nature*, due its essential aim of acquiring non-empirical knowledge of the world, which is directly incompatible with the empiricism they advocated. Stroud does not claim that the very conception of metaphysics necessarily implies it cannot succeed, although at points it certainly looks as if his position will retrospectively imply this[3]. His view is that metaphysical dissatisfaction does not lie in the conception of metaphysics but in the relation between specific metaphysical questions and fundamental features of human thought. Given this view, we would expect Stroud to offer conditions for metaphysical

[2] Stroud (2011:4): "Metaphysical reflection is meant to reveal how those thoughts and beliefs of ours are really related to the independent world they appear to be about."

[3] For example, the claims that some positive metaphysical conceptions are indispensable to human thought, and that metaphysical satisfaction requires impartial, open inquiry, combine to imply that the project of 'metaphysics from within' is unsatisfiable from the outset, due to its aim to independently assess and refine our metaphysical commitments.

satisfaction in order to then argue that they cannot be met in certain cases. And this is exactly what he does with the case of beauty[4].

He claims (2011: 10-15) that to reach a satisfactory answer to the metaphysical question of whether beauty objectively exists in the world, rather than just as a feature of human thought and responses, we would need to answer three interconnected questions:

> How is it to be established that the metaphysical answer is correct? (*ibid*:10-12)
>
> What are our actual beliefs about beauty that we are subjecting to metaphysical scrutiny? What are these beliefs like? (*ibid*: 12-13)
>
> Can we establish whether we do in fact think and respond in the ways proposed by the answer to (2), and simultaneously accept the metaphysical verdict about beauty? (*ibid*:13-14)

The first condition requires that the metaphysical verdict has epistemological support, which seems like a reasonable condition on any philosophical result. Stroud assesses some candidates for the epistemological grounds for a negative metaphysical verdict which says that beauty is *only* 'in the eye of the beholder'. A potential candidate to support the negative metaphysical verdict is the claim that a judgement of beauty is simply an expression of a personal response to an object, that is, '*that* nothing more *than the effects objects have on our minds...is relevant to the judgements we make about the beauty of things – that the "beholders'" responses themselves are enough to account for all the differences amongst different judgements of beauty.'* (*ibid*).

The claim that judgements of beauty are essentially subjective supports the claim that beauty is only a feature of our subjective conception of the world if and only if our judgements of beauty are in fact like this. Stroud's second condition therefore requires that the description of our actual conception of the world and our actual beliefs that we are testing *do in fact have the features that support for the metaphysical verdict depends on.* For example, with respect to beauty, the description of our beliefs that the subjectivist verdict depends on will be false if our judgements of beauty are in fact ascribing an objective property to objects,

[4] 'The apparently uncontroversial case of beauty seems to support at least the general feasibility of some such project.' Stroud (2011:10)

rather than expressing our personal attitude towards them. This would not undermine a subjectivist verdict about beauty, for it could be that our judgements of beauty are categorically false for attempting to ascribe an objective property which does not in fact exist, but a false description of our beliefs would undermine the support for the metaphysical verdict, and, Stroud (2011:12-13) claims, undermines the metaphysical aim of drawing a connection between our conception of the world and the world itself.

So the final condition for metaphysical satisfaction is that we are able to verify that the description of our actual beliefs is accurate *'prior to and independently of accepting any particular metaphysical verdict'* (Stroud 2011:13), i.e. whether we can verify the accuracy of the description of our beliefs and *at the same time* accept the metaphysical verdict (*ibid*:14) [5].

In summary: metaphysical satisfaction requires that we have both an accurate account of how we think about the world, and an account of how the world is, and that the former supports the latter, and that we are able to verify this with independent support. Metaphysical satisfaction can also be achieved when we find that the way we think about the world is inconsistent with an independently supported metaphysical verdict. In such a case, satisfaction requires rejecting or revising our previously held beliefs about the world. So a further condition on metaphysical satisfaction is that we are able to reject or revise the relevant beliefs if they are undermined by a metaphysical verdict (Stroud 2011:15-16).

If our metaphysical verdict says of some our beliefs about the world that they are false, then accepting the metaphysical verdict is *inconsistent* with continuing to endorse these beliefs. Here we arrive at Stroud's metaphysical *dissatisfaction* thesis. If we cannot help but continue to endorse these beliefs, then we find ourselves facing a dilemma

[5] However, Stroud's insistence that we know that the description of our beliefs that supports a metaphysical verdict is correct seems unreasonably strong to me. If we are externalistic about epistemic justification, all that matters is that the description *is* likely to be correct, rather than that we are justified in believing it to be correct. In effect, Stroud argues that metaphysical satisfaction requires that we have second order justified beliefs about our support for our metaphysical verdict, and that dissatisfaction occurs when we cannot have these second order beliefs at the same time as endorsing the verdict, or at all. These standards for metaphysical success and satisfaction seem far too high, and I shall address this in the final section.

of either accepting the metaphysical verdict at the cost of consistency, or else failing to reach a satisfying metaphysical verdict at all. Stroud claims that certain concepts are *indispensable* to human thought, such that we cannot possibly give up beliefs to the effect that the objects in question do in fact exist in the world, so we would be in an inconsistent state if we accepted a negative metaphysical verdict about them. And furthermore, the indispensability of these beliefs imply that a positive metaphysical verdict is unsatisfying since default support for it is not enough, since the question was never truly open due to the positive metaphysical conception that human thought has to presuppose.

Stroud's metaphysical dissatisfaction thesis relies on two phenomena: the inconsistency in accepting a verdict that undermines other endorsed beliefs, and the indispensability of positive beliefs about certain metaphysical objects in human thought. Importantly, he takes these two phenomena to undermine the satisfaction of the aims of metaphysical reflection. I'll dispute that claim by question what the inconsistency might be, and arguing that it is too hasty to conclude that this inconsistency necessarily undermines metaphysical satisfaction, particularly before we have a clear handle on what this inconsistency is. I will also dispute the claim that indispensability undermines the metaphysical satisfaction resulting from accepting a positive metaphysical verdict.

3. Identifying the Inconsistency

As Stroud acknowledges (2011:137), the inconsistency involved in accepting negative metaphysical verdicts about causality, modality and value which says that our conception of the world is false, is analogous to the inconsistency involved in Moorean sentences of the type 'It is raining but I do not believe it is raining'[6]. The varied examples display an apparent inconsistency between two or more propositions, but it cannot be *logical* inconsistency since the propositions do not logically contradict: their truth is compatible. In the Moorean example above, it could be true both that it really is raining, and that I do not believe it is raining. This potential has been realised on multiple occasions, for in-

[6] As discussed, for example in Moore (1944, 1993) and by Wittgenstein in *Philosophical Investigations*, part II, x (1958:190).

stance, whenever I forget how miserable a British summer can be, and am convinced it cannot be raining, though it in fact is. With respect to Stroud's cases, the inconsistency is between the negative metaphysical verdict 'x is not part of objective independent reality' and implicit commitments to the effect of the belief-state 'I believe that x is objectively functional in the world'. As before, the inconsistency is not *logical* inconsistency; there is no contradiction for it is possible them both to obtain, that is, for x not to be objectively part of the way the world is, but for me to have commitments such that I believe and act as if x *is* part of the way the world is. An analysis of this type of inconsistency could have important bearing on Stroud's thesis.

Before I get to that, I want to address some initial confusions with Stroud's inconsistency in particular. Stroud sees inconsistency in the *commitment* to a negative verdict which says that our indispensable positive metaphysical conception is false. The contents of these inconsistent beliefs are at least indirectly logically inconsistent, but since Stroud claims explicitly that endorsing negative verdicts are not instances of logical inconsistency, it is plausible to think that the inconsistency lies instead between the *belief-states* themselves, or else in the commitments, the denials and acceptances of certain beliefs[7]. It often sounds as if Stroud's thesis is that being in a 'believing a negative metaphysical verdict'-state is *incompatible* with the state we have to be in to think at all, since this state essentially involves commitments to positive metaphysical conceptions (e.g. Stroud 2011:131-137). In this sense, the 'inconsistency' would merely be *psychological* incompatibility, such that the overarching metametaphysical thesis reduces to the claim that our cognitive limitations necessarily prevent us from endorsing negative

[7] For example, Stroud (2011:137): "[C]ertain propositions P and Q and R ... necessarily, if there is any thought and experience at all then thinkers think those propositions are true. ...no thinker could consistently deny any proposition that occupies this special indispensable position in our thought. Anyone who denied any of those propositions would also accept them, since denying them involves thinking, and accepting them is required for thought. [...] Someone who both asserts that he believes that it is raining and denies that it is raining is inconsistent in asserting what he does. The two different propositions he claims to believe are not inconsistent with each other; it is possible for someone to believe that it is raining when it is not raining. Its being believed does not imply its truth. But no one who asserts or believes both that he believes that it is raining and that it is not raining could be consistent in asserting or believing what he does."

metaphysical verdicts or indeed reaching any satisfying metaphysical result in certain domains.

However, this cannot be the thesis Stroud intends to put forward, since it is demonstrably false that our unable to endorse negative metaphysical verdicts about causation, modality or value; there are a good number of philosophers, sociologists and even scientists who at the very least appear to be committed to such fundamental subjectivism. It could be argued that these people *only* appear to endorse subjective verdicts, but that this apparent commitment is either not-representative of their belief-state, or not a genuine, deep, complete or permanent commitment. Nonetheless, there are further reasons to reject the claim of psychological incompatibility. To start, it is hard to deny that people often have *logically inconsistent* beliefs. We are not logically infallible, and whether due to faulty reasoning, unconscious beliefs, or some other factor, even thoughtful, intelligent, careful agents can arrive at inconsistent beliefs[8]. And so if we are psychologically capable of being in *these* logically-inconsistent states, it would be odd if we were psychologically incapable of committing to Moorean-style inconsistencies.

In fact, in other places Stroud's allusions to hypothetical examples in which people *do* endorse negative verdicts reveals that this almost certainly isn't his claim. For instance, he discusses a hypothetical person who denies an indispensable proposition and endorses a negative metaphysical verdict.

> The inconsistency that anyone who denied an indispensable proposition would fall into would show the denier to be confused, or self-deceived, or to be acting somehow in bad faith...a negative metaphysical verdict about what are in fact indispensable beliefs could therefore seem to reveal the metaphysical status of their contents only if that indispensability is overlooked, or forgotten or suppressed. But suppressing or avoiding or even denying that indispensability would not undermine it, and so would not eliminate the inconsistency. Stroud (2011:137-138)

[8] Though not an example of psychological evidence that people have logically inconsistent beliefs, Wason (1968) experiments with a selection task which reveals a confirmation bias that demonstrates widespread errors in applying deductive rules of reasoning. It is reasonable to think that the prevalence of similar fallacies and biases implies that on occasions, incidents of faulty deductive reasoning or logical errors do result in people endorsing logically inconsistent beliefs.

We can infer from what Stroud says here that it is indeed psychologically possible for a person to endorse a negative metaphysical verdict whilst at the same time endorsing the propositions that are somehow inconsistent with it. On Stroud's view, anyone who does endorse a negative verdict necessarily also commits to the propositions inconsistent with it, for they are indispensable to any human thought at all and therefore implicitly endorsed by every human thinker. So given that the belief-state of accepting a negative metaphysical verdict is not incompatible with the unavoidable belief-state of committing to a positive metaphysical conception, the inconsistency in such instances has yet to be explained.

I turn, then, to the attempts to explain Moorean inconsistencies. The suggestion I just explored, that the inconsistency is in fact psychological incompatibility between the belief-states, could be taken as a variant of the solution which explains Moorean inconsistencies as 'unbelievable'. As Sorensen (1988:20) points out, attempting to explain Moorean inconsistencies by claiming that they cannot be believed (or similarly, claiming they cannot be asserted) just pushes the question back, since no account has been provided of what constrains believing (or asserting). A prospective account is provided, for example, by Hintikka (1962) who diagnoses Moorean inconsistencies as 'epistemically indefensible' by using doxastic logic, and then proposing indefensibility as a constraint on belief. Whilst this approach takes it as given that whatever is wrong with Moorean sentences is a violation of a constraint on belief, Stroud takes it as given that the problem in his dissatisfaction examples is precisely that they involve inconsistency, and so he does not need to claim they cannot be believed . We can potentially apply Hintikka's approach to Stroud's project by applying the doxastic logic to show that the inconsistencies in Stroud's examples are instances of epistemic indefensibility.

However, Hintikka's solution appeals to assumptions and conditions that Stroud cannot use[9]. His doxastic logic uses the rule that belief

[9] Not least because the approach is quite generally problematic for its use of rules and assumptions that many would not accept. For example, It has been shown (Linsky 1968) that this doxastic logic is too strong for entailing that the perfectly acceptable non-Moorean beliefs that someone else believes in an inconsistent proposition, is indefensible in the same way as Moorean sentences.

collects over conjunction, and the rule that a belief in a proposition implies a second-order belief that one believes in that proposition. In this way, a Moorean-style sentence becomes belief in a logical inconsistency, and Hintikka then proposes one condition of indefensibility as precisely believing a logical inconsistency. For example, the sentence 'It is raining and I believe it is not raining' becomes 'I believe that it is raining and it is not raining', which is indefensible for being a belief in the logical contradiction 'It is raining and it is not raining'.

For Stroud's cases, the inconsistency does not necessarily lie directly between two conscious beliefs which, conjoined, amount to the proposition 'x and I do not believe that x'. For there to be the kind of inconsistency Stroud worries about, the beliefs that are inconsistent with the negative metaphysical verdict need only be derivable from the indispensable positive metaphysical views, rather than included in them. In its application to Stroud's cases, Hintikka's doxastic logic would need to be adapted to extend the rule that belief in p implies a belief that one believes p, to the rule that belief in p implies belief in all the beliefs derivable from p. We then get a belief in a proposition like 'I believe that x and I believe that not x' where x is a sentence positing a positive metaphysical view of something. As before, we can then apply belief collection over the conjunction to get a belief in a logical inconsistency, which is 'indefensible' by Hintikka's logic.

Though this is how Hintikka's approach might be used to explain Stroud's notion of inconsistency, I don't think the approach is available to Stroud, due to the necessary extension of the rules. Stroud claims that the inconsistency arises even when the indispensability of the positive metaphysical beliefs are denied, that is, when the beliefs are unconscious or suppressed. Since Hintikka' approach only explains the problem with the apparently inconsistent Moorean examples when the Moorean sentence is believed, it does not help explain what is inconsistent or problematic about the situations that Stroud claims are inconsistent, but do not involve beliefs in propositions whose contents are inconsistent. In the same vein, the rule of belief-collection is inappropriate for Stroud's cases, since he claims that the inconsistency is between the distinct commitments rather than a feature a singular commitment to an inconsistent conjunction. So Hintikka's account does not help

explain what Stroud means when he says that endorsing negative metaphysical verdicts involves inconsistency.

An alternative proposal, then, is to see the inconsistency as an '*internal disharmony*' in which there is a *conflict of commitments* even when there is no conscious commitment to an inconsistency. The idea is that the actual commitments have the *potential* to lead to a direct commitment to a logical inconsistency, or *would* involve belief in a direct logical inconsistency if all the implications of the commitments were also believed[10]. If we were to accept such suggestions, we could nonetheless ask as to the severity of the problem with being *potentially* logically inconsistent, or having commitments that are otherwise structurally flawed or disharmonious. As an answer, it might be claimed that being potentially inconsistent, or structurally flawed, violates a normative requirement on us, since an 'ideal' rational agent would not be in such a state. However, a normative standard requiring rational or logical omniscience is far too strong, and even if we did not need to go this far, we might nonetheless be troubled by appeals to ideal agents in order to condemn the actual commitments of actual people as 'dissatisfying'. In general, as I shall argue in the next section, it is simply not obvious that the kind of inconsistency found in Stroud's cases implies that such cases fail to meet some normative standards.

It is hard to deny that accepting a negative metaphysical verdict which says that other beliefs we commit to are false is inconsistent and problematic in some way. But though there is clearly something to Stroud's claims, it remains unclear exactly what this inconsistency and this problem is. Though there is plenty of room left to explicate the notion of inconsistency, as it stands, it is far too hasty to conclude that the prevalence of this inconsistency undermines commitments to negative metaphysical verdicts.

---

[10] These suggestions are discussed by Sorensen (1988:36-40) e.g.: "By claiming that Bohr is committed to a direct inconsistency, we are not attributing a direct inconsistency to him. In order to arrive at the undesirable commitment, we began with Bohr's bizarre belief and deduced further beliefs. These deductions are not plausible if they intended to be psychologically accurate..." (ibid:36) and "Although incompleteability is not inconsistency, incompleteable belief-structures carry the danger of collapsing into inconsistent ones... the potential harm need never actualize." (ibid:37)

4. Overcoming Dissatisfaction

For the sake of argument, I will now just assume that the notion of inconsistency claimed to prevail in endorsements of negative metaphysical verdicts can be satisfactorily explicated. I will also grant Stroud's very arguable claim that commitments to the objective ontological status of causality, modality and value are indispensable to human thought, though I'll admit to being less than convinced on these points. My challenge is to the claim that the indispensability of certain thoughts and the inconsistency that therefore arises is so problematic as to render the particular metaphysical projects *dissatisfying*. In short, I'll argue that it is possible to satisfy our metaphysical aims in understanding modality, causality and value *in spite* of the indispensability and the inconsistency. First I'll explore an example in which a negative verdict is accepted and yet the inconsistency does not seem as dissatisfying as Stroud would have it. Second, I'll explore the potential for satisfaction in accepting *positive* metaphysical verdicts, in spite of the fact that indispensability implies that the question was not entirely open to start with.

Stroud's thesis essentially relies on the view that the inconsistency involved in accepting negative verdicts undermines the satisfaction of the aim of inquiry. Examples of dissatisfactory results seem be cases in which a person unwittingly commits to inconsistent beliefs by not recognising this inconsistency, or not recognising the indispensability of the inconsistent thoughts. But what about the metametaphysician, inclined to be reflective on the prospect of metaphysical inquiry? If she recognises that endorsing a negative metaphysical verdict involves some kind of inconsistency, or is aware that certain metaphysical commitments, though not necessarily true, are essential to human thought, might she not fare better with respect to her metaphysical ambitions?

As already cited above, Stroud (2011:137-138) claims that a negative metaphysical verdict will seem to reveal a metaphysical truth 'only if' the indispensability of the relevant inconsistent commitments is not attended to. From this, we can infer that being in an inconsistent state necessarily requires that the person does not consciously accept that he is unable to give up the commitments that are inconsistent with the negative verdict. Stroud's view seems to be that anyone who came to a negative conclusion would, taking it to contradict with a positive metaphysi-

cal conception, deny that he was committed to that positive conception, or else resolve to reject any *prior* commitment to the positive conception. In contrast, if a person *does* accept indispensability, Stroud presumably believes that they will not deny their implicit commitment to a positive conception, nor will they attempt to reject the positive conception. So Stroud's condition denies the potential for the metametaphysician to endorse a negative metaphysical verdict whilst being aware that she has indispensable positive metaphysical commitments, since the negative verdict will seem compelling *only if* the indispensable beliefs are denied or suppressed. Stroud apparently commits to the view that it is impossible for a person to recognize that some of her commitments are inevitably false, or at least, impossible for a person to believe that a commitment she cannot give up is false. I suggest that not only is this possible, it is a situation in which a person can reach a satisfying metaphysical verdict.

As an example, we can imagine the metametaphysician endorses a negative verdict about causality but recognizes that a positive metaphysical conception of causality is indispensable to any thought at all. She therefore commits to a belief that causality is not part of objective reality (negative metaphysical verdict) and a belief that she inevitably commits, to the contrary, to a positive conception of causality (acceptance of indispensability), and to a positive conception of causality (indispensable commitment). Whatever is wrong with this state is still not logical inconsistency, since again, the content of the commitments are compatible; this does not change with the addition of a conscious awareness of indispensability.

Since the negative negative verdict is not directly logically inconsistent with the indispensable positive conception, we need to represent the commitment to a negative verdict and the commitment to the indispensable positive conception as commitments of different orders, or else commitments whose implications are logically inconsistent, but whose contents are not. For example, endorsing a negative metaphysical verdict might involve the belief 'I believe that causality is not part of objective reality' without the commitment to the first order belief itself, or the indispensable positive conception might involve a belief that 'causality plays a role in the world' which does not contradict the negative verdict,

though a logical consequence of it, that there is such a thing as causality in the world', does.

If the former, the metametaphysician who is aware of indispensability could be represented as committing to:

1. I believe that causality is not part of objective reality (negative metaphysical verdict),and

2. I believe that causality is part of objective reality (acceptance of indispensability), and

3. causality is part of objective reality (indispensable positive commitment).

This state is not logically consistent, since it might be true that the metametaphysician has conflicting beliefs about causality, and that causality is part of objective reality. If it is true, then the metametaphysician would be logically inconsistent for having the contradicting beliefs she believes she has, and so believes something (not stated above) that is false. If it isn't true, then at least one of the metametaphysician's beliefs stated above is false, though she is not necessarily logically inconsistent. Perhaps this seems even more dissatisfying than Stroud's original explanation of inconsistency, but I contend that it isn't. It is possible to have a rational belief-set containing false beliefs. More controversially, it is possible to have a rational belief-set that is logically inconsistent. Sorensen (1988:24) discusses some examples, including 'propositional attitude paradoxes' (MacIntosh 1984) such as the 'sophisticated preface paradox' which says 'At least one of my first order beliefs is false'. As Sorensen (1988: 23) notes, logical consistency is just one desirable feature amongst competing desiderata for belief-states. And so if rationality is given significant weight in the aims of metaphysical inquiry, a metametaphysician who recognises that she cannot help but commit to a positive conception but nonetheless commits to a negative verdict might achieve metaphysical satisfaction in spite of the inconsistency that Stroud diagnoses, and in spite of a potential logical inconsistency. It may be objected that this is unsatisfactory precisely because the metametaphysician inevitably beliefs something false, and at the very

least, metaphysical satisfaction requires true results. Though this is a valid objection, and certainly central to Stroud's claims, I see the requirement that *all* our beliefs be true to be unreasonable and far too strong. If a negative metaphysical verdict is true, and acceptance involves rational commitments, then this seems like a metaphysically satisfying result to me, particularly if the false beliefs are the indispensable ones that the negative verdict we commit to says are false.

Finally, I think we can avoid the metaphysical dissatisfaction conclusion by denying that satisfaction requires a genuinely open inquiry, that is, an impartial starting point. Stroud claims:

> I think the indispensability of certain fundamental ways of thinking presents an unavoidable obstacle to any metaphysical satisfaction of the kind we seek. That is what stands in the way of our gaining the kind of distance we need for reaching a satisfying verdict one way or the other on the metaphysical status of those ways of thinking.
>
> For propositions whose acceptance is indispensable to any conception of a world at all, it looks as if we could never face a genuinely open question about their capturing or not capturing the way things are in the independent world.
>
> If we can never detach ourselves even temporarily from our acceptance of indispensable beliefs of those kinds, we could not subject them to the kind of scrutiny that an impartial metaphysical assessment of their relation to reality would seem to require. (Stroud 2011:140)

Whilst I agree with Stroud that the impossibility of consistently accepting a negative metaphysical verdict gives no independent support for a positive metaphysical verdict, I do not share his view that the indispensability of positive metaphysical conceptions undermines any support for a positive verdict. It is fair to say that any support we could give would not, strictly speaking, be independent of our prior commitment to a positive conception, since the thought and argument itself depends on it, if Stroud's indispensability claims are true. But this is not to say that we have no positive reason to endorse a positive verdict, nor to say that such endorsement is dissatisfying. There are many cases in philosophy where the support for a conclusion is only compelling if presupposed. 'Meta-inductive' arguments, such as No Miracles Argument in support of scientific realism, and the pessimistic meta-induction in support of

anti-realism, are such examples[11]. In these cases, the support is not independent, but this does not necessarily imply that the argument is viciously circular. It even might show that the position is at least not internally incoherent, which could well be a satisfying result.

We almost never start philosophical reflection from impartial positions, and even recognizing this rarely leads to despair about the satisfaction of our projects. Stroud imposes a stricter standard for the satisfaction of the particular metaphysical inquiry he assesses because it involves scrutinizing our own beliefs, which seems to demand impartiality on pain of circularity or inconsistency. I think the demand of impartiality, of metaphysical reflection which does not presuppose a commitment either way, is unreasonable and unsatisfiable from the outset. Since Stroud claims that it is not the very conception of metaphysical inquiry which prevents its satisfaction, in contrast to the claims of the logical positivists, impartiality cannot be an essential feature of satisfiable metaphysical projects, even when these projects involve 'metaphysics-from-within'. Though there is much more to say on this matter, in short, I claim that Stroud's conditions for metaphysical satisfaction are too strong for requiring that we are able to independently support our results. This, in turn, requires that we can verify the accuracy of our support for our conclusions without presupposing the conclusion itself, which is to hold metaphysical inquiries to a higher standard than many other philosophical pursuits.

5. Conclusion

I've here proposed three reasons why I do not take Stroud's metametaphysical thesis as conclusive. First, more needs to be said for us to understand the inconsistency that Stroud takes to be pervasive in endorsing negative metaphysical conclusions to certain inquiries. Second, Stroud has not convinced me that this inconsistency necessarily undermines the satisfaction of these metaphysical inquiries, even granting his conditions for satisfaction. Third, and finally, his conditions for satisfaction are too strong for requiring that a conclusion can be inde-

[11] See, for example, Psillos (1999) for discussions of No Miracles Arguments, and Laudan (1981) on the Pessimistic Meta-Induction.

pendently supported. Whether I give up the demand for consistent commitments, or impartial inquiries, I claim that it is reasonable to think that my metaphysical conquests *can* satisfy.

Literature

Hintikka, J. *Knowledge and Belief.* Ithaca: Cornell University Press, 1962.

Laudan, L. "A Confutation of Convergent Realism." *Philosophy of Science* 48, no. 1 (1981): 19-49.

Linsky, L. "On Interpreting Doxastic Logic." *The Journal of Philosophy* 65, no. 17 (1968): 500-502.

MacIntosh, J. J. "Some Propositional Attitude Paradoxes." *Pacific Philosophical Quarterly* 65/1 (1984): 21-25.

Moore, G. E. "Russell's Theory of Descriptions." In: *The Philosophy of Bertrand Russell*, edited by P. Schilpp, 175-225, La Salle: Open Court, 1944.

———. *Selected Writings*, edited by T. Baldwin. London: Routledge, 1993.

Psillos, S. *Scientific Realism: How Science Tracks Truth.* London: Routledge, 1999.

Sorensen, R. A. *Blindspots.* Oxford: Oxford University Press, 1988.

Stroud, B. *Engagement and Metaphysical Dissatisfaction.* Oxford: Oxford University Press, 2011.

———. *The Quest for Reality: Subjectivism and the Metaphysics of Colour.* Oxford: Oxford University Press, 2001.

Wason, P. C. "Reasoning about a Rule." *Quarterly Journal of Experimental Psychology* 20 (1968): 273–281.

Wittgenstein, L. *Philosophical Investigations*, 2nd edition, translated by G. Anscombe. Oxford: Blackwell Publishing, 1958.

Skepticism, Invulnerability, and Epistemological Dissatisfaction

Chris Ranalli
(University of Edinburgh)

How should we understand the relationship between the contents of our color, causal, modal, and evaluative beliefs on the one hand, and the color, causal, modal, and evaluative properties, on the other? According to Barry Stroud's recent book (2011), he thinks that because of how we should think about the contents of those kinds of beliefs, we should think that what he calls a "negative metaphysical verdict" on the latter is not one that we could consistently maintain. The metaphysical project aims to arrive at an improved conception of ourselves and our relation the world, no matter if that conception is positive or negative. But if Stroud is right that we cannot consistently arrive at the view that all of our causal, modal, and evaluative beliefs are systematically false, we will see that we cannot consistently reach the negative verdict. But failure to reach the negative verdict doesn't mean that we have reached the positive verdict. Stroud calls this philosophical failure "metaphysical dissatisfaction". In this paper, I argue that we can appropriate a meta-epistemological response to the problem of the external world which shares its core features with Stroud's (2000a, 2011) arguments, but which nevertheless leaves us with a distinctive kind of epistemological dissatisfaction.

1. Introductory Remarks

Since at least the publication of P. F. Strawson's *Individuals*, there has been a renewed interest in the prospects of overcoming skepticism about the external world using transcendental arguments. These arguments move from a premise about our cognitive states, such as that we think, believe, or have experiences of certain kinds, and a premise about how certain non-cognitive facts, such as that there are material objects, or that our cognitive apparatuses are reliable, are required for the truth of the cognitive premise, to a conclusion about the non-cognitive. But in his *Transcendental Arguments* and elsewhere, Barry Stroud argued that these arguments cannot be successful. However, his reasoning targeted more ambitious transcendental arguments, whose main premise connected the cognitive with the non-cognitive. We can still ask whether

or not more modest transcendental arguments, whose main premise connects cognitive states with other cognitive states, can have anti-skeptical consequences without admitting of the same problems that plagued the ambitious arguments.

In this paper, I argue that we can adapt epistemological arguments, following Davidson (1991), Stroud (1994; 2004) and a reading of Wittgenstein (1969), which aim to undermine the problem of the external world, shares its core features with modest transcendental arguments, and also parallels the kind of arguments Stroud mounts against various kinds of error theories in his recent *Engagement and Metaphysical Dissatisfaction*. The upshot of this investigation is that a better understanding of how Stroud deals with the reductionist and error theorist about color, causation, modality, and value, might bring into sharper focus another avenue of response to the problem of the external world. However, I conclude echoing Stroud's (2011), with a metaepistemological point. The point is that even a successful modest transcendental argument should leave us epistemologically dissatisfied.

2. Stroud's Metaphysical Project: An Outline

The main goal of Stroud's (2011) investigation was to discover whether or not we can arrive at significant philosophical conclusions about the nature of causation, modality, and value, after we have given an explanation of the nature of causal belief, modal belief, and evaluative belief. It is crucial here that we understand how Stroud conceives of this metaphysical project. According to Stroud, we start out with all of our beliefs, thoughts, and experiences of the world on the one hand, and we go on to ask among all of these features of human beings, which of them reflect how the world really is, and which of them distort it. Whatever survives this assessment will be what we continue to maintain, and so will receive a positive verdict. On the other hand whatever does not survive this assessment will be what we dispense with, and so will receive a negative verdict. According to Stroud, the goal of this project is to:

> [A]chieve an enhanced—a metaphysically corrected—conception of what the world is like. It is meant to tell us, contrary to the way we (perhaps uncritically) took things to be at the beginning, how things really are (Stroud 2011: 8).

But Stroud's main focus is on the prospects of a *negative outcome* of this kind of investigation. For example, we might consider a philosopher of color who maintains that, while we all have beliefs about the colors of things, nevertheless nothing in the world around us is colored.[1] What does it take to be a philosopher who maintains this view about the nature of color and color belief? Stroud considers first what it takes to have the kinds of beliefs about color that this theorist is interested in. According to Stroud, the contents of our color beliefs cannot be reduced to contents which can be understood in non-color terms. The truth of this claim rules out the contents of color beliefs from being understood in terms of the surface reflectance properties of objects causing us to have certain kinds of sensations, or from them being understood in terms of just patterns of responses to the surface reflectance properties of objects, and so on. After all, these are all properties of things which can be understood, if not in our current science of color vision and material objects, then in a completed science of color vision and material objects. This claim gives us one of Stroud's core claims:

> (Anti-Reductionism) The contents of our color beliefs are irreducible to the non-colored.

If Stroud is right that anti-reductionism is true of the nature of our color beliefs, then the next question to pursue is this: are there really colors in

[1] It's not easy to bring the full picture of what Stroud means by the "philosophical project" into view. One might compare it to Nagel's (1986) philosopher who attempts to reach a "view [of] the world from nowhere within it" (Nagel 1986: 67). Stroud expresses the "philosophical project" he's interested as follows: "If we want to find out what is really so, we have no choice but to start where we are, with what we already believe or think we know, and go on from there. We have to keep some beliefs or other, or get new ones which win out over those we used to have. But the philosophical project also requires a certain withdrawal or detachment from the things we believe. We have to be able to examine and assess them independently of our believing them or having very good reasons to believe them. The question is not simply whether they are true or whether we should continue to accept them. Even if all the beliefs we examine are the outcome of careful investigation and are as strongly supported as they can be, the question is not their epistemic credentials but the further metaphysical question of their relation to an independent reality. It is perhaps not surprising that a metaphysical enquiry into the relation between our beliefs and reality should require both engagement with and detachment from our beliefs about the world. The question is whether we can take those opposing attitudes towards the same set of beliefs. And there is the further question whether the project requires that we do so at the same time" (Stroud 2000a: 28).

the world around us? According to Stroud, his opponent finds it difficult for colors to fit into the natural world order. The core thought here is that if we aim to understand human color vision in terms of the science of color vision:

> We would see that there dose not have to be a world of coloured things to explain why we believe that there is one. Colours, therefore, would be exposed or unmasked as no part of the world as it is independently of us. They would still fall at best on the side of 'appearance', of as something 'subjective' (Stroud 2000: 76-77).

Let us focus on the last two sentences. Here, Stroud describes his opponent as someone who thinks that colors are not part of the mind-independent world: the world which is, in Bernard Williams diction, "there anyway". So we might characterize this thesis as follows:

> (Subjectivism) Colors do not exist in the mind-independent world.

This is not to claim that colors fail to exist full stop. Instead, the colors of things are something which arise due to our distinctive responses to things, or are otherwise properties of certain kinds of mental states like visual experiences. In either case, color properties count as being mind-dependent: properties whose existence depends upon the existence of minded creatures.

At this juncture, Stroud asks a good question. He asks whether or not the reductionist about the content of color beliefs can also be a subjectivist about colors. His answer is 'no', because he thinks that in order for us to *ascribe* color beliefs to ourselves and others, we must also *believe* that objects are colored. In his words, the ascription of the belief that some things are colored is *indispensible* to ascribing others with color beliefs. They are indispensible in the sense a *necessary condition* of our ascribing color beliefs to people is our actually believing that objects are colored.[2] This gives us the claim that:

> (Stroud's Color Thesis) A necessary condition of believing that *S has color beliefs* is believing that some things are colored in the world around us.

[2] Here's how Stroud expresses it: "no one could abandon all beliefs about the colours of things and still understand the colour terms" (2000a:168).

Stroud's master argument, then, amounts to this. If anti-reductionism and Stroud's Thesis are true of the nature of our color beliefs, then the philosopher of color cannot consistently arrive at subjectivism about color. In more general terms, what we discover about the nature of color beliefs constrains what we can learn about the metaphysics of color. According to Stroud, what we cannot learn is that there are no colored things independently of us. But it's not that we cannot learn that there are no colored things independently of us because there *are* colored things independently of us. Stroud doesn't think that the failure to reach the negative conclusion about the colors of things implies that we have successfully reached the positive conclusion. Instead, we will be faced with the dissatisfying conclusion that one of two possible outcomes of our metaphysical project is unreachable, but not because the other possible outcome has been reached.

We can abstract from the details of the color case, and express Stroud's general line of argument as follows:

> Template Stroudian Argument:
> (Belief Claim) S believes that *a*, *b*, *c*, and so on, are *F*.

For example, S might believe that certain events cause others, or that people do things for *reasons*, or that some propositions *couldn't have been false*. But now suppose that there is no causation in the independent world, no reasons in the independent world, and so on. In general:

> (Negative Verdict) There are no *F*s.

At this stage, we register Stroud's core thesis that indispensible to believing that *a* causes *b* is believing in material objects, or that indispensible to believing that people have color beliefs is believing that things are colored, or that indispensible to having beliefs at all is believing things *for reasons*. In general:

> (Indispensability) A necessary condition of believing that *a*, *b*, *c*, and so on, are *G*, is believing that *a*, *b*, *c*, and so on, are *F*.
>
> Therefore,
>
> (Invulnerability) S cannot consistently maintain that *a*, *b*, *c*, and so on, are *G* while accepting the *Negative Verdict*: that there are no *F*s

Of course, an important and difficult part of this project is filling out the Belief Claim. This is what we might call *Stage 1* of the Metaphysical Project: to give the correct account of the contents of our beliefs in the relevant domain. Stroud wants to make sure that he has given the correct account of the contents of our beliefs in the relevant domain. Consider causation. He thinks that to believe that *a* caused *b* is to believe that there is some *necessary connection* between *a* and *b*, irreducible to counter-factual analysis and dispositional analysis. But if we then go to the world to see whether or not anything ever satisfies the contents of those beliefs, it seems as if nothing could. Call this *Stage 2* of the Metaphysical Project. What Stroud finds interesting is whether we can *consistently maintain* the verdicts of Stage 1 of the Metaphysical Project with the verdicts of Stage 2 of the Metaphysical Project. If we cannot, then we have failed our Metaphysical Project. Following Stroud, call this *Metaphysical Dissatisfaction.*

If Stroud is right that substantive areas of our thought, such as our causal, modal, and evaluative thoughts, are *indispensible* to our thinking at all, or to our thinking in certain fundamental ways, then it seems he would also be right in thinking that the outcome of our general assessment of them cannot be negative. The sense in which the outcome cannot be negative is *not* the strong modal sense in which it cannot be true that people fail to have these kinds of thoughts. After all, it is possible for us not to exist, as we once did, and so possible for us not to have thoughts at all. Instead, Stroud's point seems to be that it is not possible for us have thoughts of those kinds, *consistent with* accepting a negative conclusion about whether or not things in the mind-independent world answer to those concepts which in part compose those thoughts.

In this way, a negative verdict about the possibility of correctness for these kinds of thoughts cannot be reached; so subjecting them to this kind of metaphysical assessment, where a negative outcome is seen to be one of two possible improvements, should strike us as dissatisfying. It's dissatisfying not because it cannot be rewarding. *It can be*. It can be rewarding to learn that certain ways of thinking are so fundamental to thinking anything at all, or thinking in certain ways, that subjecting them to metaphysical assessment turns out to be pointless. It can be rewarding to learn that certain arguments for unappealing metaphysical

conclusions turn out to be fallacious. Rather, Stroud's point seems to be that what promises to leave us dissatisfied is that one of the most general goals of metaphysics is to reach a refined and improved conception of the world, no matter how magnificent or bleak that conception of the world turns out to be. But if arriving at an improved conception of the world turns out to be impossible for us, how *could* we be satisfied with that? His point here is not to mount this argument as an argument against, say, an error theorist about colors, who believes that if our color concepts are ever satisfied in the world independently of us, they are satisfied by categorical color properties of objects, but since there are no such 'spooky' properties in the world around us, it follows that our beliefs about the colors of things, that some carpets are red, or that lemons are yellow, say, are systematically false. Instead his point here is to mount this argument as an argument that is supposed to enhance our *understanding of philosophical questions*, *philosophical explanations*, as opposed to, say, 'ordinary questions', 'scientific questions', 'religious questions', 'ordinary explanations', 'scientific explanations', and so on. So the target of his arguments, it seems, are meta-philosophical, rather than strictly first-order, philosophical concerns. The meta-philosophical point is that certain philosophical questions, such as 'are there any colors independent of anyone?', are not questions we can consistently give a negative answer to, but not because the answer to 'are there any colors independently of anyone?' is '*no*'. Instead, we are *unable* to *consistently believe* that, say, there are no colors independently of anyone, because, surprisingly, metaphysical reflection on the nature of color and color belief revealed that "no one could abandon all beliefs about the colours of things and still understand the colour terms" (Stroud 2000*a*:168).

Now that we have examined the structure and aims of Stroud's(2000*a*, 2011) meta-philosophical arguments I want to ask whether or not this sort of meta-philosophical argument can extend over into the epistemological domain, where one of the primary concerns is over the possibility of knowledge of the external world. One motivation here is that prima facie similarities between the two philosophical projects might not turn out to be superficial similarities, but authentic. If this is right, then we will be furnished with a novel meta-philosophical argument, not against skepticism about the external world, but against

the prospects of our being able to arrive at skepticism about the external world as the improved epistemological conception of ourselves and our relation to the world, the kind we might seek as epistemologist who want to understand the limits and extent of our knowledge.

3. The Epistemological vs. the Metaphysical Project

On the face of it, there is an important difference between the metaphysical projects and the epistemological project. One of the core differences between the problem of explaining how our color, causal, modal, and evaluative judgements could be true independently of us and the problem of explaining how our knowledge of the world is possible is that the latter, unlike the former, seems to present us with a certain kind of *paradox*.

Consider Stroud's famous discussion of the problem of the external world in his (Chapter 1, 1984). There, Stroud argued that the problem of the external world seems to be grounded in "platitudes we all accept" (Stroud 1984: 82). But if the problem is grounded in grounded in "platitude we all accept", then this presents us with a distinctive metaepistemological problem: how is it possible to avoid epistemological dissatisfaction when it comes to responding to the problem of the external world? After all, if Stroud is right that what makes a skeptical answer to that problem look plausible is that the argument in favor of skepticism about the external world trades on nothing more than platitudes, then a response to that argument would have to reject platitudes. But what kind of epistemological satisfaction does that promise? Call this *Stroud's Paradox.* Any resolution to Stroud's Paradox seems to leave us with a kind of epistemological dissatisfaction. On the one hand, skepticism about the external world does not reflect how we think of ourselves and our epistemic relation to the world. After all, we think of ourselves as knowing about the world around us. On the other hand, if the argument in support of that thesis trades on mere platitudes, then in order to avoid skepticism, we would have to revise one or more of the platitudes that we all would otherwise accept. But this kind of revisionism in our epistemological thinking is at the same time

what we seek to avoid in our philosophical theorizing, but also what we cannot avoid if we are to oppose skepticism about the external world.

The problem here for our purposes is not Stroud's Paradox, but that it is hard to see how our response to skepticism about the external world could parallel Stroud's response to the subjectivist and reductionist about colors, causation, value, and so on, if the former, but not the latter, is grounded in mere platitudes. To help sharpen the problem, consider the dissatisfaction which Stroud argues accompanies the project of explaining how our color, causal, modal, and evaluative beliefs could be independent of our responses to the world. On his view, it arises out of the relationship between a view about the content of those kinds of beliefs, and what we think is one of two possible outcomes of the distinctive philosophical project of explaining how those kinds of beliefs could ever be true. One outcome is the negative outcome where none of them could be true. Of course, this might be a disappointing result because we perhaps want to think that our modal, evaluative, and causal beliefs, for example, could be true of the world around us. But what is not disappointing as a philosopher is that we could arrive at the truth of how things are in the world around: that it is possible to arrive at a conception of the world where none of our modal, causal, color, or evaluative discourse represents the world as it is. The problem here, then, is appears to be a stark contrast between the two projects, and the source of the philosophical dissatisfaction each project seems to recommend. Call this the *Contrast Problem*.

One problem our account faces, then, is the Contrast Problem. How should it be overcome? In what follows, I argue that the Contrast Problem is in fact not a genuine problem for giving a response to skepticism about the external world which parallels Stroud's response to the subjectivist and the reductionist. On this view, while there seems to be a problem for giving some such response, appearances are misleading.

What we should examine now is the problem of the external world, and the distinctive epistemological project of explaining how our knowledge of the world around us is possible. Concerning the problem of the external world, Stroud wrote at the beginning of his *Significance of Philosophical Scepticism*, that:

> [T]he problem has no solution; or rather that the only answer to the question as it is meant to be understood is that we can know nothing about the world around us (Stroud 1984: 1).

Notice that this quotation presents us with an intellectual challenge, because Stroud is not himself a skeptic. So we should ask what he means if not that skepticism about the external world is true. One interesting point to highlight about this quotation is that Stroud makes it look like once we understand the problem as it is "meant to be understood", we too will find that the answer is that we can know nothing about the world around us. But how do reconcile a commitment to this claim with the commitment that skepticism about the external world is false?

Perhaps the place to start is this: how should we understand the epistemological problem of the external world? After all, it might be that the problem has "no solution" because a solution would show how we can or do have the kind of knowledge that skepticism claims we cannot have, *given* the distinctive challenge to that knowledge which it presents us with. So, perhaps Stroud thinks that there is "no solution" to the problem because no proposed solution could show how our perceptual knowledge of the world around us is possible, *given* the distinctive obstacles which make that knowledge look impossible. Compare with Zeno's challenge: how is motion in general possible *if* it is also true that we must traverse an infinite distance in order to move at all? Zeno's challenge is to show how movement is possible *given* those 'facts' about motion that he raises. The challenge is not to show how movement is possible full stop. If that were right, we could just run around a track, or open a door.

Moreover, notice that understanding the question like this reveals that our question 'how is knowledge of the external world possible?' is elliptical for 'how is knowledge of the external world possible *given* such-and-such obstacle which makes it look impossible?' On this reading, the question is a conditional question. It does not ask how something is possible full stop. Instead, it is, at least at the semantic level, what Cassam (2007) has called an *obstacle-dependent question*: a question which is raised at all because something makes the target of the question *look impossible.*

But what is it that makes knowledge of the world look impossible? As it stands in the literature, the answer to this question is 'all sorts of things'.[3] Of course, we could inquire into which of these obstacles is more basic than the others, in the sense that it might be that one of them is indispensible to raising the problem of the extenral world, and sustaining the skeptical answer to that problem, but this would bring us too far afield. Instead, I want to bracket this important question, and just focus on what Stroud has taken the obstacle to our knowledge of the world to be. Let us therefore consider a core statement of the problem as Stroud understands it:

> The Cartesian problem of our knowledge of world therefore becomes: how can we know anything about the world around us on the basis if the senses give us only what Descartes says they give us [my emphasis] (Stroud 1984: 12).

---

[3] One obstacle is thought to be grounded in the *Closure Principle* and the contraints it puts on propositional knowledge. The Closure Principle can be expressed as follows: if S knows that p, and S knows that p implies q, and deduces q from p while retaining her knowledge that p throughout, S thereby comes to know that q. The problem the Closure Principle raises is two-fold. On the one hand, if we plug in certain skeptical scenarios, such as that we are brains-in-vats, the Closure Principle puts us in a position to reason as follows: I know that having hands entails that I'm not a handless brain-in-a-vat. So, from the Closure Principle, it follows that if I know that I have hands, I know that I'm not a brain-in-a-vat. But how do I know that I'm not a brain-in-a-vat? At this juncture, Nozick (1981) and Dretske (2005, 1970) argued that we cannot know that we are not brains-in-vats, because if we were brains in vats, we would nevertheless believe that we weren't. Still, it apparently strikes some as just *obvious* that we couldn't know some such proposition (Cf. DeRose 1995: 2). But this is where the second problem arises. The Closure Principle seems to be no more than a theoretical expression of the basic idea that deduction can be a means of our extending our knowledge (Cf. Williamson 2000). So the Closure Principle presents us with the two-sided problem of having to show how it is possible to know that we are brains-in-vats even though it's 'obvious' that we couldn't know some such proposition, or showing that Closure Principle is false, even though it seems like nothing more than an expression of the platitude that deduction can be a means of our extending our knowledge to the known implications of what we know. Another obstacle is about sense perception and its relation to empirical knowledge. Since at least the publication of Descartes' *Meditations,* most philosophers have thought that it is possible to that we have all of the sensory experiences that we have, even if everything we believe about the world is false. Call this the *Underdetermination* of sensory experience. For responses to this obstacle in particular, see Burge (1998), Bruekner (1994, 2010), McDowell (2008, 1995), Pritchard (2012), Pritchard and Ranalli (*forthcoming*), Stroud (2009), and Cassam (2007). However, there are some philosophers—Contextualists—who think that the obstacle is grounded in a misrepresentation of our linguistic practices and their relationship to our standardards for knowldege-attribution. See Unger (1975, 1984) and DeRose (1995). For a survey of the recent work on radical skepticism, see Pritchard (2002).

So what, according to Descartes, do the senses give us?

> What we gain through the senses is on Descartes' view only information that is compatible with our dreaming things about the world around us and not knowing anything about the world by means of the senses. The Cartesian argument presents a challenge to our knowledge, and the problem of our knowledge of the external world is to show how that challenge can be met (Stroud 1984: 13).

So, as Stroud presents the problem at least, the skeptical challenge to our knowledge of the world is this:

> (Skeptical Challenge) How is knowledge of the external world possible if it is also true that our knowledge of the world comes to us through our senses, but we could have all of the sensory experiences that we have even if we knew nothing about the world around us?[4]

Notice here that the skeptical challenge is not framed in terms of an adeversial dispute between the 'Skeptic' and the 'Anti-Skeptic'. If this were the challenge, what would be required for an epistemologically satisfying response to the challenge would be different. After all, if the challenge is framed in terms of a dispute between us and 'Skeptics', then one goal would be to *convince* the skeptic to be one of us: to believe what we believe. Call this the *Convince-the-Skeptic Challenge*. As several philosophers have argued, however, this challenge appears to end in favor of the Skeptic.[5] Indeed, as Wright (1991) has pointed out, if

[4] Stroud frames the obstacle in different terms from his (1984) to his (1994). In the firse case, the obstacle is that it is possible to have all of the experiences that we have even if we are all dreaming. In the second case, the obstacle is that it is possible to have all of the experiences that we have, even if everything we believe about the world is false. However, note that these obstacles all give expression to the same core idea: that there is a gap between our sensory experience and the world around us, such that no amount of sensory experience is ever *sufficient* for knowledge of the world around us, contra the intuition that sensory experience sometimes seems to just give us knowledge of the world without any cognitive effort—without any reasoning or reflective thinking in play.

[5] Prime examples where this case is made include Stroud (1984, 1989), Pritchard (2005, 2012), Bryne (2004), Pryor (2000), Wright (1991), Williams (1991), and Rorty (1979). It seems like one of the core arguments of Stroud's (1984) was to show that the Convince-the-Skeptic Challenge must end in favor of the skeptic. There, Stroud argued that Austin, Moore, Kant, and Quine all failed on this front, in part because they were engaging in the Convince-the-Skeptic Challenge. But given Stroud's argument of his (1984), that skepticism is a paradox grounded in platitudes, he also seemed right in thinking that responses within the Convince-the-Skeptic Challenge would end in failure. For a discussion on how one might intergate a transcendental response along diagnostic lines,

Stroud's thesis about the problem of the external world is right, a sound argument against skepticism makes the problem even more acute because now we are furnished with two appealing arguments for opposing conclusions, which just "intensifies the embarrasment" against the 'Anti-Skeptic' (Wright 1991: 89).

Under the influence of Stroud (1984) and Williams (1991), most philosophers have now abandoned the Convince-the-Skeptic Challenge, and have instead embarked on the challenge of showing how the problem of the external world is far less of paradox than Stroud made it seem. On this view, the problem of the external world, and its attandent skeptical argument, do not trade on mere "platitudes that we would all accept", platitudes which are grounded in our pre-theoretical thinking about knowledge and sense-perception " (Stroud 1984: 82). Call this the *Diagnostic Challenge*. A successful response to the Diagnostic Challenge, it seems, would allow for epistemological satisfaction, since the goal would then be to expose where in the skeptical argument there are more than just platitudes in operation. But there would be no dissatisfaction in unmasking these principles as far less than platitudinous, since then we would see that we were not presented with the problem of the external world qua paradox. We would instead be in a position to see that the problem of the external world is more of a paradox along the lines of Skolem's or Cantor's paradox, where rejecting an assumption amounts to no more than problems for our theories, rather than the Sorites and the Liar paradox, where rejecting an assumption seems to have ramifcations for more than just our theories, but also the correctness of our use of ordinary concepts.

To see this, consider the following expression of the *Skeptical Paradox*:

> (1) I don't know that I am not a brain-in-a-vat, deceived by an evil demon, and so on.
>
> (2) If I don't know that I am not a brain-in-a-vat, deceived by an evil demon, and so on, then I don't know that there are tables, trees, computers, and so on.

---

as opposed to Convince-the-Skeptic lines, and how this bears on the arguments from Stroud, Williams, Davidson, and Rorty, see Pritchard and Ranalli (*forthcoming*).

(3) I know that there are tables, trees, computers, and so on.

What makes this paradox look like it is grounded in "platitudes we all accept"? Consider first premise (2). This premise can be supported along the following lines. Intuitively, if one knows, say, that they are in Washington Square park, New York City, then they know that they aren't in Patriarch's Ponds park in Moscow, Russia. Generalizing, if one knows that *p*, and one knows that *p* implies *q*, then they know that *q*.[6] This is one expression of the principle that knowledge is closed under known-entailment. One might think that this principle just gives theoretical expression to a core principle that operates in our assessment and attribution of knowledge in everyday life. After all, deduction seems to be one way of extending our knowledge to the known consequences of what we already know. The Closure Principle just articulates this intuition. And since I know that having hands, or there being trees, and so on, implies that I'm not a brain-in-a-vat, or deceived into thinking that there are hands, trees, and so on, then from Closure it follows that if I know that I have hands, I know that I'm not a brain-in-a-vat. But how *could* I know that I'm not a brain-in-a-vat? Of course, I don't often consider these kinds of far-fetched possibilities. But that I don't consider them just registers an epistemic failing on my part: a failure to consider all the possibilities relevant to whether or not I know about the world around me. And once such possibilities are raised, it looks like it *is* part of our everyday criteria for assessing and making knowledge attributions that we withdraw the claim that we know that we aren't victims of those types of scenarios, just as a lawyer might get a witness to withdraw their claim to seeing the suspect, after other possibilities have been raised.[7] Moreover, if any sensory experience I have is logically neutral with

[6] In order to avoid obvious counter-examples to K(S, p) and K(S, [p → q]) *entails* K(S, q) one the basis of the failure of the analogue principle for belief, B(S, p) & B(S, [p → q]) *entails* B(S, q), because S can fail to believe that q and because knowledge entails belief, K(S, p) → B(S, p). So, one needs to modify the closure principle so that if K(S, p) and K(S, [p → q]), then if S deduces q from p, whilst retaining their belief that p and q throughout, S thereby comes to know that q. This called 'Single Premise Closure'; for a discussion and defense, see: Hawthorne (2005).

[7] Cf. Stroud's 'Cleveland' example from his Chapter 1 'The Problem of the External World', in his (1984).

respect to the world around me, how could any sensory experience alone, or any amount of sensory experience, put me in a position to *know* that those skeptical possibilities do not obtain? But I after all *do* know about the world around me. So now we are furnished with a set of mutually inconsistent propositions, all of which are individually plausible, and look as if they are grounded in nothing more that common sense.

Thus far, we have seen what look like important points of departure between the epistemological project and the metaphysical project. The main point is that the former, unlike the latter, seems to arise out of the goal of trying to resolve a distinctive kind of *paradox*—indeed, a paradox which is supposedly "grounded in platitudes we all accept" (Stroud 1984: 82). But there are also important points of contact between the two projects:

> (1) Both can be framed as an attempt to resolve certain philosophical 'How-Possible?' questions.
>
> (2) Both seek to arrive at an improved conception of ourselves and our epistemic relation to the world, no matter if this conception is positive or negative.
>
> (3) Both cases seem to lure the threat of philosophical dissatisfaction of some kind.

In the next section, I aim to strengthen the comparison between the two projects, and so overcome the Contrast Problem. In order to do this, I will first examine the role of transcendental arguments can take as responses to skepticism about the external world, but later show how one can appropriate a transcendental argument from Davidsonian content externalism which shares its core features with Stroud's response to the error theorists. The goal here is to show how the two kinds of problems admit of similar responses. In section 4, I argue that there is another potential source of philosophical dissatisfaction other than the kind I reviewed earlier. While the former kind of philosophical dissatisfaction arises from attempting to respond to the problem of the external world if it is also true that it is grounded in platitudes, this kind of philosophical dissatisfaction resembles, in its structure, the kind Stroud identified arises from responding to the various forms error theories and subjectivism.

4. Transcendental Arguments: Strategies and Problems

4.1 Ambitious Transcendental Arguments

Since at least the publication of P. F. Strawson (1959) *Individuals*, there has been a renewed interest in the prospects of transcendental arguments against various kinds of skepticism. The aim of Strawson's transcendental reasoning, like Kant's before him, was to reveal certain substantive facts about the world around us on the basis of a priori reasoning about the conditions required for certain substantive mental phenomena. For example, in his (1959), Strawson argued that we have a certain cognitive capacity, a capacity to re-identify what he called "particulars" in experience—these are the referents of our demonstrative expressions like 'this' or 'that', and can be things like ordinary material objects—and that a necessary condition of having this capacity is the continued existence of particulars unperceived. From these two premises a substantive fact about the world around us follows: that particulars continue to exist unperceived.

But in his (1968), Stroud famously argued that transcendental reasoning of this kind must end in failure. According to Stroud, there are two basic problems with transcendental reasoning. The first problem is that the conclusions of these kinds of arguments do not follow unless some kind of idealism or verification principle about meaning is true. But both of these doctrines are contentious. And what is more, either of them alone are *sufficient* to block the skepticism the transcendental reasoning was meant to avoid. Thus transcendental reasoning is either ineffectual or superfluous. *Ineffectual* because if we dispense with either idealism or the verification principle, these arguments lose their anti-skeptical consequences. *Superfluous* because if we keep either idealism or the verification principle, the transcendental reasoning adds nothing to the anti-skeptical consequences of these positions. Call this *Stroud's Dilemma*.[8]

[8] For an excellent collection, and explanation of, traditional arguments in epistemology, see Stern (1999).

4.2 Stroud's Dilemma

What supports Stroud's Dilemma? In order to understand this claim, we should look at a concrete case of transcendental reasoning. Consider Strawson again. We can represent the structure of his reasoning as follows:

> (P1) We can re-identify particulars in sensory experience.
>
> (P2) A necessary condition of the capacity to re-identify particulars in sensory experience is the existence of particulars unperceived.
>
> Therefore,
>
> (C) Particulars exist unperceived.

The problem Stroud raised is that, if *S* is a proposition whose truth is a necessary condition for the possibility of thought, experience, belief, and so on, "the skeptic can always very plausibly insist that it is enough to make [thought, experience, belief, ...] possible if we believe that *S* is true, or if it looks for all the world as it is, but that *S* needn't be actually true" (Stroud 1968: 225). Call this the *Hedging Move*. After the Hedging Move is made, the transcendental claim will be:

> (P2*) A necessary condition of the capacity to re-identify particulars in sensory experience is that we *believe* in the existence of particulars unperceived.

But now the validity of the argument turns on establishing a *necessary connection* between our *believing* in the existence of particulars unperceived *and* the existence of particulars unperceived. In general, the criticism is that if the Hedging Move is legitimate, then the validity of transcendental arguments turns on establishing a necessary condition between *believing that S is true* and *S being true*. Otherwise, the arguments are invalid.

At this juncture, one would need to appeal to some principle in order to bridge the gap between our believing that S is true and S being true. Call this the *Bridge of Necessity move*. On this point, Stroud says that:

> It would seem that we must find, and cross, a bridge of necessity from the one to the other. That would be a truly remarkable feat, and some convincing

> explanation would surely be needed of how the whole thing is possible (Stroud 2000: 158-159).

But is it possible to close this gap unless we appeal to some kind of idealism or verificationism, which deflates the gap between *our believing* what we believe and *the truth* of what we believe? Stroud thinks the answer to this question is 'no'. But then this renders the transcendental argument superfluous, since these doctrines alone have anti-skeptical consequences.

Now, one might dispute either the Hedging Move, the Bridge of Necessity Move, the claim that idealism or verificationism are required to cross that bridge, or that these doctrines alone have anti-skeptical consequences. I want to bracket all of these responses, and instead concentrate on the Hedging Move. We might think that even if the Hedging Move is legitimate, this does *not* prevent the transcendental argument from having anti-skeptical consequences. For example, if it's true that our believing certain non-psychological propositions are necessary conditions of us realizing certain psychological states, we might wonder whether more 'modest' transcendental claims of this kind can protect us from skepticism.

4.3 Davidsonian Externalism and Modest Transcendental Arguments

Recall that the Hedging Move weakens the core premise of the transcendental argument from:

> (Ambitious Claim) S could not have certain cognitive states (reidentification of particulars), unless certain non-cognitive states obtained (particular exist unperceived).

To:

> (Modest Claim) S could not have certain cognitive states (reidentification of particulars), unless S believed that certain non-cognitive states obtained (particular exist unperceived).

The proponent of a modest transcendental argument thinks that weakening the core premise in this way does not diminish the anti-skeptical force of a transcendental argument. But how is this possible? Unlike their ambitious cousins, modest transcendental arguments move

from premises about how certain psychological states or attitudes require certain other psychological states or attitudes. But the claims are not supposed to be trivial. The claims are supposed to link two kinds of seemingly different types of psychological states or attitudes. For example, Stroud argued in his (2000*b*) that a necessary condition of attributing beliefs about colors to people is believing that certain things are colored. Of course, this should sound surprising. After all, I can attribute beliefs about ghosts to people without believing that there are ghosts. So what makes beliefs about the colors of things so special? Indeed, this is one of the goals of modest transcendental arguments: to explain what *is* so special about certain kinds of psychological states or attitudes.

But even if the claims are supposed to be non-trivial, the question still stands: how *could* anything anti-skeptical follow from propositions about how certain psychological states or attitudes are necessary conditions of certain others? Let us consider the externalism about belief content found in the work of Donald Davidson. Davidson's brand of content externalism hinges on his notion of *triangulation*. Triangulation is a causal relationship, holding between at least two persons and one object, where each person 'triangulates' on, or identifies and is receptive to, a common cause of their belief or utterances. Crucially, Davidson seems to think that triangulation is required for us to have beliefs and utterances with determinate contents at all. For example, he says that:

> Without this sharing of reactions to common stimuli, thought and speech would have no particular content that is, no content at all. It takes two points of view to give a location to the cause of a thought, and thus to define its content. We may think of it as a form of triangulation: each of two people is reacting differentially to sensory stimuli streaming in from a certain direction. If we project the incoming lines outward, their intersection is the common cause. If the two people now note each other's reactions [...] each can correlate these observed reactions with his or her stimuli from the world. The common cause can now determine the contents of an utterance and a thought. The triangle which gives content to thought and speech is complete. But it takes two to triangulate. Two, or, of course, more. (Davidson 1991: 159-60)

Since, on Davidson's view, triangulation is required for determinate content, this makes his view a kind of content externalism, because the causal nexus between the persons and the common cause is of course external to what the person can access in introspection, or otherwise is

external to the intrinsic properties of those persons. Of course, all that that passage revealed is that Davidson endorses a certain kind of externalism about the content of belief and speech. But this alone does not tell us what it is about some such view which lends it to anti-skeptical argument. On this front, we should consider another statement of Davidson's view:

> It should now be clear what ensures that our view of the world is, in its plainest features, largely correct. The reason is that the stimuli that cause our most basic verbal responses also determine what those verbal responses mean, and the content of the beliefs that accompany them. The nature of correct interpretation guarantees both that a large number of our plainest beliefs are true, and that the nature of these beliefs is known to others (Davidson 1991: 160-161).

From this passage, we can extract the following core statement of Davidson's externalism. The first reading of his view concerns the truth of our beliefs:

> (Non-Epistemic) Davidsonian Externalism: A necessary condition of having and attributing beliefs with determinate contents to ourselves and others is that some of those attributed beliefs are.

But the passage also permits an epistemological reading:

> (Epistemic) Davidsonian Externalism: A necessary condition of having and attributing beliefs with determinate contents to ourselves and others is that we are in a position to know the causes of some of those beliefs, and so in a position to know that some of them are true.

The latter entails the former, and it is hard to see how Davidson would allow the first version of the thesis to stand without also endorsing the second version. After all, if in order to attribute beliefs with determinate contents to ourselves and others, how could I fail to be in a position to know the causes of those beliefs, if the content of those beliefs are what we triangulate on in our environment? Instead, it looks more like, if we were unable to know the causes of what we attribute to ourselves and others, then Davidsonian externalism would end up entailing that we could not know what we believe.

This issue takes us to an important point of interest. We saw that Stroud raised what looked like knock-down arguments against the anti-skeptical prospects of ambitious transcendental arguments. But it does not look like Davidsonian content externalism lends itself to the same

general criticism. After all, how could Davidsonian content externalism presuppose idealism or verificationism, when it is just a thesis about the relationship between having thoughts with certain contents and the world around us? The implication is that if we have thoughts of a certain kind, then we have had causal-contact with the world. But what is it about that conditional which is verificationist or idealistic? Moreover, we can accept that Stroud is right that even an ambitious transcendental argument from Davidsonian content externalism would be susceptible to the Hedging Move. But the response would be that the Hedging Move counts as a good objection to the anti-skeptical prospects of an ambitious transcendental argument insofar as the aim of these arguments are to convince the skeptic: to engage in the Convince-the-Skeptics Challenge. But if the aim is instead to reveal an interesting relationship between thought and the world, irrespective of the Convince-the-Skeptic Challenge, it's not obvious what would be wrong with some such argument. The thought here is that one can mount an transcendental argument for more reasons than just anti-skeptical reasons; so we shouldn't see arguments against the anti-skeptical prospects of transcendental arguments as arguments against the interestingness and usefulness of transcendental arguments.

Indeed, the point can be pressed futher: what prevents one from giving an ambitious transcendental argument from Davidsonian Externalism? Davidsonian externalism would have anti-skeptical consequences in conjunction with knowledge of the contents of our thoughts. The thought here is that, provided that I know what I believe, and provided that I know that certain external conditions, such as that a certain external object caused me to have the belief that I have, are necessary conditions of having a belief with that content, I am now in a position to know, by deduction, about those external causes. This would be a straightforward argument from self-knowledge and content externalism to claims about the external world. Indeed, since the premises move from propositions about belief and the external conditions required for having beliefs, to conclusion about the external world, it is unclear what prevents us from motivating an ambitious transcendental argument from externalism about content.

As it turns out, in certain places it looks as if Stroud rejects the core statement of Davidsonian externalism. Davidsonian externalism seems

to have the straight-forward implication that from a thinkers having a thought with a certain content, it follows that a certain kind of object is present in the thinker's environment. But according to Stroud:

> What calls into question the validity of the last step of would-be transcendental arguments from the way we think to the way things are is the apparently simple logical observation that something's being so does not follow from its being thought or believed to be so. Something's being so does not follow from everyone in the world's believing it to be so, from everyone's fully reasonably believing it, even from every reasonable person's being completely unable to avoid believing it (Stroud 2000: 166-167).

Here, it looks like Stroud is just registering the logical intuition that propositions about belief and thought do not entail propositions about the world around us, and vice-versa. Whether or not one ought to have the intuition is a good question. After all, intuitions are defeasible, and it seems like content externalism gives us principled reasons for thinking that, while propositions about the world do not entail propositions about thought and belief, it is possible for some propositions about thought and belief to entail propositions about the world, when the content of those states are externalist in character. For example, if content externalism is true, then if I am thinking that water is wet, and if the concept WATER has externalist individuation conditions, it *does* follow that water exists, even though from the fact that water exists, it does not follow that someone is thinking about water. Of course, this is all compatible with thinking that if the direction of entailment went from *water exists* to *someone is thinking about water* is valid, then we are compromising our view that the world is mind-independent because it would mean that there just *could not be* a situation in which p obtains but no thinker is *thinking that p.* Still, externalism should make the direction from thought to world look less clear. If the existence of certain kinds of things are necessary, not for our thinking at all, but for our thinking about those kinds of things in particular, then what is wrong with thinking that propositions describing that we are thinking about them entail that those things are there to be thought about?

So, it is not obvious what reason Stroud might have for avoiding the ambitious transcendental argument from content externalism, modulo the intuition we registered above. Of course, even if he were to endorse some such argument, he would still face what looks like formidable

obstacles to that argument.[9] But those reasons would not be among the general reasons he levied against the success of ambitious transcendental arguments.

Let us now move on the core question of this section: how might Davidsonian content externalism have anti-skeptical consequences if it figures in a modest transcendental argument? Recall that we saw the challenge to our knowledge of the external world could be put as follows:

> (Skeptical Challenge) It is possible that, while I believe propositions about the world on the basis of experiences which are compatible with their falsity, I cannot know that those propositions are true.

The epistemological reading of Davidsonian Externalism permits a straightforward response to this challenge. After all, the problem of the external world is the problem of explaining how knowledge of the world around us is possible. The question arises because certain obstacles like the one above make knowledge of the world look impossible. So what we want to explain is how it is possible that, while we have all these beliefs about the world around us, we can ever know that these beliefs are true. One answer to this question is skeptical. It says that we can't ever know that any of our beliefs about the world are true. So this is supposed to be one of two possible outcomes. It's the negative outcome. The positive outcome maintains that we can know that some of our beliefs about the world around us are true. Davidsonian externalism implies that the negative outcome cannot be consistently reached. On the first non-epistemic reading, Davidsonian externalism implies that the truth of some of our belief about the world is a necessary condition of our

[9] The core challenge on this score is that transcendental arguments from content argument seem to license unacceptable forms of reasoning. The first premise is a premise expressing some piece of self-knowledge (e.g. that I am thinking about water); the second premise is a premise expressing what we know, on the basis of a priori argument, to be required for that kind of self-knowledge (e.g. that content externalism implies that if I am thinking about water, then water exists), to conclusions about the external world (e.g. that water exists). But, so the objection goes, how can I deduce that certain propositions about the external world are true, from non-empirical propositions? The upshot is that priviledged self-knowledge looks incompatible with content externalism. This is called the McKinsey Paradox, after Michael McKinsey's original statement of the problem. See McKinsey (1991); for responses see Wright (2008), Davies (1998); for recent work on McKinsey's Paradox, see Kallestrup (2010).

maintaining them. So if the problem is to assess all of our beliefs about the world around us, and ask whether and how knowing that any of them are true is possible, our going on to do what the problem of the external world asks us to do—identify a wide portion of our beliefs—guarantees, by externalist lights, that we find that some of them are true. On the second, epistemic reading, Davidsonian externalism implies that the possibility of knowing that some of our beliefs about the world around us are true is a necessary condition of attributing beliefs with determinate content to ourselves and others. So we could not consistently arrive at the conclusion that, while we have all sorts of beliefs about the world, it is nevertheless not possible that we know any of them.

In order to raise the problem of the external world, and engage with it, I have to assess some beliefs about the world around me. This implies that I at least know what I believe about the world around me. From this claim and Davidsonian externalism, it follows that I am in a position to know that some of those beliefs are true. I provided a tentative argument for the view that it is unclear how Stroud's general critique of ambitious transcendental arguments applies to this case, because it looks like Davidsonian externalism provides that bridge we needed in order to move from premises about our cognition to conclusions about the non-cognitive world, without presupposing a problematic idealism or verificationism.

However, we can also mount a modest transcendental argument from Davidsonian externalism. The first things we need to do is hedge the Davidsonian claim:

> (Epistemic) Davidsonian Externalism*: A necessary condition of having and attributing beliefs with determinate contents to ourselves and others is that believe we are in a position to know the causes of some of those beliefs, and so in a position to know that some of them are true.

With this hedged definition of Davidsonian externalism in play, we are now in a position to provide the following argument:

> Modest Davidsonian Argument:
>
> (Belief Claim) S believes, and attributes to others, that there is water, trees, people, planets, and so on, for all beliefs about the external world.

(Negative Verdict) S cannot know whether or not propositions about the external world are true.

(Modest Davidsonian Indispensability) A necessary condition of the truth of the Belief Claim, is that we believe that we can know whether or not some of those propositions are true.

Therefore,

(Modest Davidsonian Invulnerability): I cannot consistently maintain that while I believe that there is water, trees, people, planets, and so on, that nevertheless I cannot know whether or not those beliefs are true.

Three points to notice about this argument. The first point is that this argument is not targeting skeptical arguments which seek to show that knowledge of the world is impossible. If it were targeting those kinds of arguments, we would have slipped back into the Convince-the-Skeptic Challenge, which we presumed ends in favor of would-be skeptics. Instead, the argument is targeting the epistemological project itself—the goal of explaining how it is possible for us to know about the world around us. The conclusion of the Modest Davidsonian argument is not that the epistemological project is meaningless or uintelligible. Instead, the conclusion of the Modest Davidsonian argumetn is that it is not one that we can *consistently engage in.*

The second point is that the arguments against the success of ambitious transcendental arguments are ineffectual here because our core transcendental claim is modest rather than ambitious. Of course, this does not mean that the argument is unproblematic. Instead, it just means that the problems with the argument are not grounded in general problems with transcendental arguments.

The third point and final point is that Negative Verdict—that skepticism about the external world is true—need not be established in order for the Modest Davidsonian argument to be effective. We can see straight-off that the negative outcome of the epistemological project is that it's not possible to know that our beliefs about the world around us are true. After we've identified that as one outcome of the epistemological project, we can go on to ask whether it is possible for us to reach that negative conclusion rather than the positive conclusion. Prima facie at least, it seems like it should be possible. After all, we can

ask similar epistemological questions, such as whether it is possible to know that some of our beliefs about ghosts and hauntings are true. The negative answer would be that we cannot. So what makes the general epistemlogical question about our knowledge of the world around us any different?[10]

Notice too that this makes it look like first two premises of the Modest Davidsonian argument are secure. The soundness of the argument therefore seems to turn on whether or not it's true that if we have beliefs about the world around us, we must also believe that we can know that some of those beliefs are true. To be sure, this view isn't obvious. But neither is any interesting philosophical thesis.

Finally, what is the anti-skeptical important, if any, of the Modest Davidsonian argument? Recall that if the aim of the epistemological project is to arrive at an improved conception of the epistemological relations between ourselves and the world around us, the Davidsonian argument implies that the negative outcome of this project is unreachable. But it is unreachable not because the argument has shown that skepticism about the external world is false. Instead, it is unreachable because our epistemologist must, at the same time, have some beliefs about the world in order to subject them to epistemological assessment, but also, according to the core Davidsonian claim, believe that it is possible to know that some of those propositions are true. The latter is seen as required for the former, and this renders the negative skeptical outcome unreachable.

To sharpen the issue, it seems the problem facing the epistemological project arises from a tension between the following three ideas:

[10] Williams (1991) thinks there *is* a significant difference. To ask what epistemological class our 'knowledge of the external world' is in, is to ask whether there is a common factor between all of the things we claim to know in that class from which a general explation can be given. Prima facie, knowledge of the external world is knowledge of tables, chairs, people, ice-cream, churches, phones, and so on. It is a kind of knowledge which would be about the objects, properties, and events we know about in the world around us. But Williams is skeptical that this forms a genuine epistemological class from which a general epistemological question should be raised.

(Epistemological Goal) The goal of arriving at an improved epistemological conception of our relation to the world around us, no matter whether or not this conception is positive or negative.

(Belief Possession) Epistemologists engaging in the (Epistemological Goal) must have some beliefs about the world around them in order to engage in the problem of explaining how our beliefs about the world around us could ever amount to knowledge.

(Modest Davidsonian Indispensability) If S has beliefs about the world around them, then S believes that it is possible for us to know whether or not some of those beliefs are true.

On the one hand, Belief Possession and Modest Davidsonian Indispensability make it look impossible to achieve the Epistemological Goal because it implies that one of two possible outcomes cannot be reached, but not because the positive verdict is false. But since Belief Possession seems to be secure, this result might make us look at whether or not the Epistemological Goal is intelligible, or the truth of a modest form Davidsonian externalism. Perhaps because the intelligibility of the Epistemological Goal and the centrality of Belief Possession to that goal are non-negotiable, this might put pressure on Davidsonian externalism.

5. Indispensability & Invulnerability

So far, we've see that if Davidsonian externalism is true, one can mount a certain kind of meta-epistemological argument against the possibility of reaching a negative epistemological conclusion to a core question of epistemology: 'how is it possible to know that any of our beliefs about the world around us are true?'. What I want to explore in this section is whether there are any other philosophical theses which might make our target epistemological question one that cannot consistently be given a negative answer, but not because it is impossible for us to know anything about the world around us.

Let us work backwards. We can understand the core premise of the transcendental argument in Stroud's terms as follows:

The truth of Q is *indispensible* to truth P.

In that:

> S could not have (P) certain cognitive states (reidentification of particulars) unless (Q) certain non-cognitive states obtained (particular exist unperceived).

In broad terms, we can call this kind of indispensability *Truth Indispensability*. Thus the core features of the more ambitious transcendental arguments are that certain non-cognitive states are truth-indispensable to certain cognitive states. Of course, there are different classes of what might called truth-indispensible propositions. For example, a proposition might be truth-indispensible to our having any cognitive states at all. This would be an extreme case. But less extreme cases are also possible, such as with Strawson's transcendental argument. In that case, the obtaining of certain non-cognitive states of affairs were truth-indispensible to a particular kind of cognitive-state, rather than them all.

This is where modest transcendental arguments diverge from the ambitious arguments. Modest transcendental arguments weaken the core premise of the transcendental argument as follows:

> It is not possible that S have (P) certain cognitive states (e.g., reidentification of particulars), unless (Q) S *believed* that certain non-cognitive states obtained (e.g., particular exist unperceived).

Call this *Conceptual Indispensability*, because it maintains that certain cognitive states, such as belief, are indispensible to certain others, in the sense that the having of one kind of cognitive state or attitude is a necessary condition of having another kind of cognitive state or attitude. The modest Davidsonian transcendental argument is an example of an argument which argues from conceptual indispensability to invulnerability: the invulnerability of certain beliefs of ours to certain kinds of challenges.

Are there other kinds of Indispensability which fit the format of modest transcendental arguments? Consider the following: S could not engage in practice P unless S *believed* that certain non-cognitive states obtain. Call this *Methodological Indispensability*. A belief in a proposition can be methodologically indispensible if it satisfies the following principle:

> (Methodological Indispensability) If S is engaging in practice P, then S cannot consistently reject Q if believing Q is a necessary condition of engaging in practice P.

Methodological Indispensability implies Methodological Invulnerability:

> (Methodological Invulnerability) If S is engaging in practice P, then S cannot consistently reject Q.

After all, if it's true that our *believing* Q is a necessary condition of our even engaging in practice P, then how *could* we consistently engage in that practice, and nevertheless *reject* Q? It would not be possible. Of course, it would be possible to engage in the practice, and reject Q. But we would just be *inconsistent*.

An example of an argument from methodological indispensability to methodological invulnerability can be found, I think, in Wittgenstein's notes publised as *On Certainty*. A core idea that Wittgenstein discusses is tthat of a 'hinge' commitment. One expression of the idea can be found in the following passage:

> [...] the questions that we raise and our doubts depend upon the fact that some propositions are exempt from doubt, are as it were like hinges on which those turn. That is to say, it belongs to the logic of our scientific investigations that certain things are in deed not doubted. But it isn't that the situation is like this: We just can't investigate everything, and for that reason we are forced to rest content with assumption. If I want the door to turn, the hinges must stay put. (OC 11 §§341-3)

How should we understand the metaphor of the 'hinges' which belong to what Wittgenstein called "the logic of our scientific investigations"? One view is that these 'hinges' are propositions. But these propositions are not just true propositions that are practically necessary or *needed* in order to engage in our scientific investigations. But neither are these propositions necessary truths. Instead, these propositions have a certain peculiar *epistemological role* in our epistemological practices such as assessment, evaluation, assertion, proving, doubting, conjecturing, arguing, and so on. It is these propositions which *make intelligible* those epistemological practices in the sense that those propositions *operate for*, but not *within*, those epistemological practices. For example, in §125 Wittgenstein considers a kind of skeptical attack on the claim that he has hands:

> If a blind man were to ask me "Have you got two hands?" I should not make sure by looking. If I were to have any doubt of it, then I don't know why I should trust my eyes. For why shouldn't I test my eyes by looking to find out whether I see my two hands? *What* is to be tested by *what*? (OC §125).

His response sharpens the idea that doubting the proposition that he has hands, in this case, would constrain what he could appeal to in order to alleviate his doubts. If he appealed to his senses, why wouldn't whatever reason he had for doubting his hands not also extend to his senses?

We might put the point like this. Certain propositions are methodological necessities, in the sense that their truth, while not *necessary truths*, are necessary for engaging in the kinds of epistemological practices we in fact engage in. So the propositions could have been false. Moreover, their truth is not *guaranteed by* our engaging in the epistemological practices we engage in. We could engage in those epistemological practices, and those propositions could nevertheless be false. But what we could not consistently do is *accept* that our epistemological practices are legitimate while *denying* those propositions. This furnishes us with another example of an anti-skeptical argument which fits our Stroudian template. Plausible hinge-propositions might include the proposition that our senses are reliable, or that an external world exists. These propositions form part of the set of beliefs that we are subjecting to epistemological assessment. But if Wittgenstein is right, then these propositions *must* remain unassessed, on pain of making unintelligible the epistemological assessment we are engaging in. In this case, our *belief* in the their truth is required to make intelligible the epistemological assessment. But the problem of the external world is supposed to be a problem which arises from an assessment of *all* of what we believe about the world around us. Insofar as the propositions that there is an external world or that our senses are reliable are hinge propositions, then, those propositions cannot enter into that assessment. In this fashion, those propositions are *invulnerable* to a negative outcome of that assessment. But those propositions are also *indispensible* to engaging in that assessment. So, if this Wittgensteinian argument is sound, it follows that we cannot *consistently raise* the problem of the external world, and so cannot be consistently presented with that problem.

6. Epistemological Dissatisfaction

Philosophical dissatisfaction represents a kind of philosophical failure. Describing the precise nature of this failure and how it arises is an important question, but outside the scope of what I want to achieve here. The kind I am focusing on here is *goal-directed*. We have a certain philosophical goal in mind, and it turns out that certain compelling arguments reveal that that goal cannot be achieved, leading to philosophical dissatisfaction.

Let us return to the third point of §2. Against the Contrast Problem, the point was that both the epistemological project and the metaphysical project seems to lure philosophical dissatisfaction of some kind. But the Contrast Problem highlighted how the philosophical dissatisfaction that the epistemological project lures is grounded in a certain conception of the problem of the external world. This conception of the problem of the external world is a conception where the problem is a paradox grounded in "platitudes we would all accept" (Stroud 1984). The problem then is that, a response to the problem of the external world would have to reject one or more platitudes. But the rejection of mere platitudes is philosophically unsatisfying. So there seems to be a stark disanology between the epistemological project and the metaphysical project. In the epistemolocial case, philosophical dissatisfaction arises, if it does, because we have to reject platitudes, and so revise deep-seated intuitions, concepts, or principles. But in the metaphysical case, philosophical dissatisfaction arises, if it does, from discovering that we are unable to consistently reach a certain outcome of an intelligble, perhaps even desirable, philosophical goal.

Recall that is Davidsonian Externalism is true, it does not follow that we have knowledge of the world around us. It does not even follow that it is not possible that we know nothing about the world around us. All that follows is that we could not consistently maintain and so arrive at what is supposed to be one of two possible outcomes of our general epistemological assessment—the negative, skeptical outcome. Notice that this conclusion is reminiscent of the conclusions that Stroud reached in his (2011), strengthing the case against the Contrast Problem. What I want to focus on now is one of the consequences Stroud urged

followed from those reflections: that we would be dissatisfied with this outcome.

It is important to note that Stroud is not claiming that his arguments, and perhaps arguments of the same general kind, will leave us dissatisfied. The arguments might be good arguments and in this sense should leave us satisfied. Instead, what Stroud is claiming is that, given what we were aiming for in asking our general philosophical question—either a positive or a negative verdict—in order to arrive at an improved conception of the world, no matter if that conception left us with a magnificent world or a bleak world, the argument will leave us dissatisfied with respect to that philosophical question. But if Stroud's arguments are successful, then we cannot reach the negative verdict: we cannot reach what was supposed to be one of two possible outcomes of our philosophical assessment. But what is perhaps worse is that failure to reach one of the outcomes does not *entail* that we have reached the other possible outcome: the positive verdict. Instead, we would have found that while we cannot reach the negative verdict, failure to reach that verdict is alone not a reason for believing that we have arrived at the positive verdict.

The next step is appropriating this conclusion for the epistemological case. In the epistemological case, we saw that Davidsonian externalism seemed to have anti-skeptical consequences, even when we applied it in a modest transcendental argument. One consequence that was brought out is how a negative answer to the problem of the external world could not be reached if Davidsonian externalism is true. This would mean that one of two possible outcomes of a general assessment of our beliefs about the external world could not be reached. We could not arrive at the negative verdict of that assessment if Davidsonian externalism is true. But still, does that mean we have now explained how knowledge of the world is possible? *No.* For one, that we cannot consistently arrive the negative skeptical verdict is not explained, in this case, by its falsity. In this way, we haven't even shown that knowledge of the world around us is possible, if the modest transcendental argument from Davidsonian externalism is sound. Likewise, Davidsonian externalism alone appears to be silent on whether knowledge of the world around us is possible. All it gives us is a

conditional claim; so nothing about whether the kind of knowledge it quantifies over is possible follows from accepting that conditional.

Moreover, Wittgenstein's thesis about hinge-propositions gave us the materials for constructing a modest transcendental argument against skepticism. It showed us that, if propositions like *there is an external world* are hinge-propositions, then these propositions are invulnerable to the negative outcome of the epistemological assessment of all of our beliefs about the world, because that proposition makes some such assessment intelligible. But that that proposition is included in the assessment is indispensible to it as well. In this fashion, the argument from Wittgenstein's hinge-propositions shows us that we cannot consistently engage in *that* kind of assessment, and so cannot consistently engage with the problem of the external world. It is here, I think, where if Wittgenstein is right about there being a class of propositions that are hinge-propositions, a certain kind of epistemological dissatisfaction ensues. The dissatisfaction shows up in the intuitive thought that the problem of the external world is at least intelligible to us, and something that we could engage with as rational agents. But the Wittgensteinian point is that that is not the case.[11]

If philosophical satisfaction and dissatisfaction are receptive to the success or failure to realize our philosophical goals, then it looks like, if our modest transcendental arguments are sound, the goal of understanding and so explaining how our knowledge of the world is possible in general is problematic. It is problematic not because we have shown that the positive outcome of pursuing that goal cannot be reached. Rather, it was shown that if our modest transcendental arguments are sound, the negative outcome of pursuing that goal cannot be reached. But then one of two possible outcomes—one of two possible improvements in our conception of ourselves and our epistemological relation to the wider-world—will have been revealed as unreachable. This, I think, should leave us dissatisfied.[12]

---

[11] For different treatments of anti-skeptical import of Wittgenstein's 'hinge-proposition' strategy, see Pritchard (*forthcoming*) and Wright (2004). For a review of interpretations of Wittgenstein's response to skepticism, see Pritchard (2011).

[12] I'd like to thank Duncan Pritchard, Allan Hazlett, and Barry Stroud for helpful discussions on topics related to this paper. I'd also like to give a special thanks to the

Literature

Bruekner, A. "Modest transcendental arguments", *Philosophical Perspectives* 10 (1996): 285–300.

———. "The Structure of the Skeptical Argument", *Philosophy and Phenomenological Research* 54 (1994).

———. "~K~SK". In: *Essays on Skepticism.* Oxford: Oxford University Press, 2010.

———. "Fallibilism, Underdetermination, and Skepticism." In: *Essays on Skepticism.* Oxford: Oxford University Press, 2010.

Burge, T. "Cartesian Error and the Objectivity of Perception." In: *Subject, Thought and Context*, edited by P. Pettit and J. McDowell, 117-136, Oxford: Oxford University Press, 1986.

Cassam, Q. *The Possibility of Knowledge.* Oxford: Oxford University Press, 2007.

Davidson, D. "Three Varieties of Knowledge." In: *A. J. Ayer Memorial Essays*, edited by A. P. Griffiths, 153-166, Cambridge: Cambridge University Press, 1991 (reprinted in Davidson 2001, chapter 13).

Davies, M., "Externalism and Armchair Knowledge." In: *New Essays on the A Priori*, edited by P. Boghossian and C. Peacocke, 384-415, Oxford: Clarendon 2000.

Descartes, R. *The Philosophical Works of Descartes.* Translated by E. S. Haldane and G. R. T. Ross, New York: Cambridge University Press, 1986.

DeRose, K. "Solving the Skeptical Problem." *Philosophical Review* 104 (1995): 1-52.

Dretske, F. "Epistemic Operators." *Journal of Philosophy* 67 (1970): 1007–1023.

———. "Is Knowledge Closed Under Known Entailment? The Case Against Closure." In: *Contemporary Debates in Epistemology*, edited by Matthias Steup & Ernest Sosa. Oxford: Blackwell, 2005.

Hawthorne, J. "The Case for Closure" In: *Contemporary Debates in Epistemology*, edited by Matthias Steup & Ernest Sosa, Oxford: Blackwell, 2005.

Kant, I. *Critique of Pure Reason*, translated by Norman Kemp Smith, New York: St Martin's Press, 1965.

Kallestrup , K. "Recent Work on McKinsey's Paradox." *Analysis* 71, no. 1 (2010): 157-171.

McDowell, J. "Knowledge and the Internal", *Philosophy and Phenomenological Research* 55 (1995): 877-893.

———. "The Disjunctive Conception of Experience as Material for a Transcendental Argument." In: *Disjunctivism: Perception, Action, Knowledge*, edited by A. Haddock & F. Macpherson, Chapter 16, Oxford: Oxford University Press, 2008.

McKinsey, M. "Anti-Individualism and Privileged Access." *Analysis* 51 (1991): 9–16.

Nagel, T. *The View from Nowhere.* Oxford: Oxford University Press, 1986.

Nozick, R. *Philosophical Explanations.* Oxford: Oxford University Press, 1981.

---

organizers and participants of the August 2013 Bamberg summer school in 'Metaphysics or Modernity?' with Professors Barry Stroud, Jorge Gracia, Saul Kripke, and Richard King.

Pritchard, D. H. "Recent Work on Radical Skepticism", *American Philosophical Quarterly* 39 (2002): 215-257.

———. *Epistemic Luck*, Oxford: Oxford University Press, 2005.

———. "Wittgenstein on Scepticism." In: *The Oxford Handbook of Wittgenstein*, edited by O. Kuusela & M. McGinn, Oxford: Oxford University Press, 2011.

———. *Epistemological Disjunctivism*. Oxford: Oxford University Press, 2012.

———. "Wittgenstein on Hinges and Radical Scepticism in On Certainty." In: *Blackwell Companion to Wittgenstein*, edited by H.-J. Glock & J. Hyman, Blackwell, forthcoming.

Pryor, J. "The skeptic and the dogmatist", *Noûs* 34 (2000): 517–49.

Ranalli, C. and Duncan Pritchard. "Rorty, Williams, and Davidson: Skepticism and Metaepistemology". Humanities" In: *Special Issue 'The Legacy of Richard Rorty'*, edited by Neil Gasgoine, forthcoming.

————. "Skepticism and Epistemological Disjunctivism", *Continuum Companion to Scepticism*, forthcoming.

Rorty, R. *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press, 1979.

———, ed. *Transcendental arguments: Problems and prospects*. Oxford: Oxford University Press, 1999.

Strawson, P. F. *Individuals: An essay in descriptive metaphysics*. London: Methuen, 1959.

———. *Individuals: An Essay in Descriptive Metaphysics*. New York: Doubleday, 1959.

Stroud, B. "Transcendental Arguments", *Journal of Philosophy* 65 (1968): 241-256.

———. *The Significance of Philosophical Scepticism*, Oxford: Oxford University Press, 1984.

———. "Kantian Argument, Conceptual Capacities, and Invulnerability." In: *Kant and Contemporary Epistemology*, edited by P. Parini, 231–51, Dordrect: Kluwer, 1994.

———. *The Quest for Reality: Subjectivism and the Metaphysics of Colour.* Oxford: Oxford University Press, 2000a.

———. *Understanding human knowledge: Philosophical Essays*. Oxford: Oxford University Press, 2000b.

———. "Scepticism and the Senses." *European Journal of Philosophy* 17, no. 4 (2009): 559-570.

———. *Engagement and Metaphysical Dissatisfaction: Modality and Value.* Oxford: Oxford University Press, 2011a.

———. *Philosophers Past and Present: Selected Essays.* Oxford: Oxford University Press, 2011b.

Unger, P. *Ignorance: A Case for Scepticism.* Oxford: Clarendon Press, 1975.

———. *Philosophical Relativity.* Oxford: Basil Blackwell, 1984.

Wittgenstein, L. *On Certainty*, edited by G. E. M. Anscombe & G. H. von Wright, translated by D. Paul & G. E. M. Anscombe, Oxford: Basil Blackwell, 1969.

Williams, M. *Unnatural Doubts: Epistemological Realism and the Basis of Scepticism*, Oxford: Blackwell, 1991.

Williamson, T. *Knowledge and Its Limits.* Oxford: Oxford University Press, 2000.

Wright, C. J. G. "Scepticism and Dreaming: Imploding the Demon." *Mind* 397 (1991), 87-115.

———. "Wittgensteinian Certainties." In: *Wittgenstein and* Scepticism, edited by Denis McManus. Routledge, 2004.

———. "McKinsey One More Time" In: *Self-Knowledge*, edited by Anthony E. Hatzimoysis. Oxford: Oxford University Press, 2008.

A Trip to the Metaphysical Jungle – How Kripke's Intuitions Revived Aristotelian Essentialism

Sebastian Krebs
(University of Bamberg)

Besides his substantial influences on other fields of philosophy, Saul Kripke is famous for smoothing the way for a new type of Aristotelian essentialism. However, Kripke's comments on essentialism are utterly vague and are built entirely on fundamental intuitions about the use of language and necessity *de re* in modal logic. He famously disproved Willard Van Orman Quine who, a few decades earlier, had banned necessity *de re* into the metaphysical jungle of Aristotelian essentialism – a jungle that a true empiricist must not enter. But not only Kripke's refutation of Quine, but also his own essentialism is based on intuition. Kripke thereby overcomes an anti-essentialist dogma that was established by Kant in his *Critique of Pure Reason*. But because of its vagueness, one might well call Kripke's intuitive essentialism a 'metaphysical jungle'. Nevertheless, a trip to Kripke's jungle is a promising milestone on the way to a refreshed Aristotelian metaphysics.

1. Travel arrangements: The rejection of Quine's anti-essentialism

In *Two dogmas of empiricism*, Willard Van Orman Quine rejects the traditional distinction between synthetic and analytic truths. Among others Immanuel Kant suggested this distinction in his *Critique of Pure Reason*. Quine criticizes it as a relic from metaphysical dogmas which a pure empiricism (as Quine wants to pursue it) has to overcome:

> [O]ne is tempted to suppose in general that the truth of a statement is somehow analyzable into a linguistic component and a factual component. Given this supposition, it next seems reasonable that in some statements the factual component should be null; and these are the analytic statements. But, for all its a priori reasonableness, a boundary between analytic and synthetic statements simply has not been drawn. That there is such a distinction to be drawn at all is an unempirical dogma of empiricists, a metaphysical article of faith.[1]

[1] Willard Van Orman Quine, "Two Dogmas of Empiricism," in *From a Logical Point of View. 9 Logico-philosophical Essays* (Cambridge: Harvard University Press, 1961), 36-37.

As a consequence, Quine gives his empiricism a pragmatic turn. There is no longer any necessary truth, but only some 'web of beliefs' which can be more coherent or less coherent. This is the reason why Quine claimed that the search for an object's essential properties would be in vain. For him, the external world is accessible via descriptions only: If there is necessity in logic at all, it must be *de dicto*, since this type of necessity is reducible to a semantic predicate. Therefore, Quine accepts necessity *de dicto* at least "for the sake of argument"[2], while he in fact rejects a modal calculus as a whole – at least he did so in his early papers on modal logic. Quine's idea of a pure descriptive necessity is best illustrated by his famous example of a cycling mathematician, respectively, a mathematical cyclist:

> Mathematicians may conceivably be said to be necessarily rational and not necessarily two-legged; and cyclists necessarily two-legged and not necessarily rational. But what of an individual who counts among his eccentricities both mathematics and cycling? Is this concrete individual necessarily rational and contingently two-legged or vice versa? Just insofar as we are talking referentially of the object, with no special bias toward a background grouping of mathematicians as against cyclists or vice versa, there is no semblance of sense in rating some of his attributes as necessary and others as contingent.[3]

Quine regards the cycling mathematician, taken as a person, neither as rational nor as two-legged. Only after one has considered the descriptions 'mathematician' and 'cyclist', does he become necessarily rational, respectively, two-legged. Kripke summarizes this position as follows:

> Whether a *particular* necessarily or contingently has a certain property depends on the way it's described.[4]

Kripke also showed that Quine's approach (which might well be called 'anti-essentialist empiricism') does require a descriptive account of reference, as suggested by Gottlob Frege and Bertrand Russell. Therefore, it is counter-intuitive. He illustrates this counter-intuitivism

---

[2] Christopher Hughes, *Kripke. Names, Necessity, and Identity* (Oxford: Oxford University Press, 2004), 81-82.
[3] Willard Van Orman Quine, *Word and Object* (Cambridge: MIT Press, 1960), 199.
[4] Saul Kripke: *Naming and Necessity* (Cambridge: Harvard University Press, 1980), 40 (original emphasis).

in *Naming and Necessity* with an example from the U.S. presidential election in 1968:

> Suppose that someone said, pointing to Nixon, 'That's the guy who might have lost'. Someone else says 'Oh no, if you describe him as 'Nixon', then he might have lost; but, of course, describing him as the winner, then it is not true that he might have lost'. Now which one is being the philosopher, here, the unintuitive man? It seems to me obviously to be the second.[5]

A similar reasoning applies to Quine's cycling mathematician: If he is described as a 'mathematician', he is necessarily rational. If he is described as a 'cyclist', he is necessarily two-legged. But it would be absurd to answer the question whether a certain person, for example Smith, is necessarily two-legged in the following way: 'If you describe Smith as a cyclist, he is necessarily two-legged. But if you describe him as a mathematician, then he is not.' Therefore, Quine's anti-essentialist understanding of modality must be rejected because it presupposes a counter-intuitive way of language use.

To stick to the jungle metaphor, we might regard Quine as the worried father who thought that our metaphysical adventure of exploring essentialism is far too costly, dangerous or philosophically pointless. With Kripke, we could say that it is our job to convince Quine of the philosophical benefits of such a trip by appealing to some fundamental intuitions – which is what I will try to do in the following parts of my paper.

2. Waiting for departure: A plea for intuition-grounded philosophy

Of course, intuition-grounded argumentation causes doubts among many philosophers. The main objection could be summarized as follows: Intuitions are purely subjective and therefore random – and such subjectivity cannot lead to philosophical insight. However, Kripke gives, first and foremost, a semantic analysis of language. This means, he reflects on what people do, in fact, express when they utter a certain sentence. However, a good semantic analysis goes hand in hand with the intuitions which language users have about their own language. For that

[5] Kripke, *Naming and Necessity*, 10.

reason, we can well reject any analysis which does not appeal to our intuitions.

To put it more theoretically, we might say that intuitions lead to a 'reflective equilibrium' as John Rawls calls the relating balance of intuitions and theory in moral philosophy.[6] This basically means that a counter-intuitive theory is not a good theory since it does not apply to most, and definitely not to all, cases of language use. Hence, Kripke gives a well-founded argument to reject Quine's anti-essentialist understanding of necessity.

However, it seems that Kripke regards intuitions as the unique way which leads to philosophical evidence:

> Of course, some philosophers think that something's having intuitive content is very inconclusive evidence in favor of it. I think it is very heavy evidence in favor of anything, myself. I really don't know, in a way, what more conclusive evidence one can have about anything, ultimately speaking.[7]

Consequently, not only Kripke's refutation of Quine is intuition-grounded, but also his own account of essentialism – if one wants to call the "better picture"[8] Kripke draws in *Naming and Necessity* an account or a theory at all. Anyway, Christopher Hughes claims correctly that Kripke's essentialism neither refers to Aristotle nor to Thomas Aquinas (only to mention two famous essentialists in history), but rather to "the man on the Clapham omnibus"[9]. For this stereotype bus passenger, it is intuitively evident that individuals have some properties which are contingent and some others which are necessary. According to this, Richard Nixon, for example, has the contingent property of being elected as U.S. president in 1968. But he has the necessary property of being human – and not an aardvark or any other non-human creature.[10] Kripke's intuitions do not only apply to individuals, but also to natural kinds. For example, 'being H_2O' is a necessary property of water, while 'being liquid' is not. I will come back to this example later.

---

[6] Cf. John Rawls, *A Theory of Justice* (Cambridge: Harvard University Press, 1971).
[7] Kripke, *Naming and Necessity*, 42.
[8] Ibid., 94.
[9] Hughes, *Kripke*, 84.
[10] Cf. Robert Stalnaker, "Possible Worlds Semantics: Philosophical Foundations," in *Saul Kripke,* ed. Alan Berger (Cambridge: Cambridge University Press, 2011), 100-115.

By such an intuitive reasoning, Kripke overcomes an anti-essentialist dogma in philosophy which was established by Immanuel Kant in the *Critique of Pure Reason* and which heavily influenced Quine and other 'early analytic philosophers' like Bertrand Russell or Ludwig Wittgenstein. However, Kripke's idea of essentialism remains utterly vague, since he – in his characteristic way of presenting philosophic ideas – only presents a few examples of essential properties and does not give any sufficient principles or general statements on his account. Because of this vagueness, one is well-justified in calling Kripke's intuitive essentialism a 'metaphysical jungle' – of course, without the negative connotations Quine had in mind by using this metaphor. On the contrary, a trip to Kripke's jungle is very promising for contemporary philosophy – not only in terms of logic and philosophy of language, but also as a milestone on the way to a refreshed Aristotelian metaphysics. Therefore, I will, in the next part of this paper, sketch some general remarks on Kripke's intuitive examples which might be seen as the entering step into this adventurous jungle.

3. Entering the Jungle: General remarks on Kripke's essentialism

So far, I have pointed out that Kripke's intuitions lead him to a metaphysical position which is, in general, called essentialism or sometimes even Aristotelian essentialism. But what is exactly meant by these terms? Here is Quine's definition:

> Aristotelian essentialism. This is the doctrine that some of the attributes of a thing (quite independently of the language in which the thing is referred to, if at all) may be essential to the thing, and others accidental. E.g., a man, or talking animal, or featherless biped (for they are in fact all the same things), is essentially rational and accidentally two-legged and talkative, not merely qua man but qua itself.[11]

This definition is quite adequate to describe Kripke's approach. Hughes characterizes this essentialism therefore as 'moderate', which means that it stands between hypo-essentialism and hyper-essentialism (cf. Hughes 2004:108-110). On the one hand, hypo-essentialism refers to a

[11] Quine, Willard Van Orman, "Three Grades of Modal Involvement," in *The Ways of Paradox and other Essays* (New York: Random House, 1966), 173-174.

theory that claims that only trivial properties are necessary, as for example 'being self-identical', 'being red or not red', 'being warm if being hot' etc. – i.e. properties all individuals have. On the other hand, hyper-essentialism claims that each individual has all of its properties necessarily – a view one might well ascribe to Gottfried Wilhelm Leibniz. However, neither hypo-essentialism nor hyper-essentialism is what Kripke has in mind – his essentialism is moderate and therefore in accordance with the given definition by Quine. It seems that Kripke takes 'essential' and 'necessary' as synonyms – even though the Aristotelian tradition defines essential properties as properties which are non-trivial and still necessary. I will not distinguish between them either, even though I focus only on non-trivial necessary properties (as Kripke obviously does as well).

Kripke mentions a simple principle which gives a certain hint in order to find out which properties are necessary and which are not: the necessity of origin. He introduces it in a footnote of Naming and Necessity as follows:

> If a material object has its origin from a certain hunk of matter, it could not have had its origin in any other matter.[12]

Kripke comes to this principle by help of two intuitive examples: his wooden table and Queen Elizabeth II:

> We could conceivably discover that, contrary to what we now think, this table is indeed made of ice from the river. But let us suppose that it is not. Then, though we can imagine making a table out of another block of wood or even from ice, identical in appearance with this one, and though we could have put it in this very position in the room, it seems to me that this is not to imagine this table as made of wood or ice, but rather it is to imagine another table, resembling this one in all external details, made of another block of wood, or even of ice.[13]

> [C]ould the Queen – could this woman herself – have been born of different parents from the parents from whom she actually came? [...] [C]an we imagine a situation in which it would have happened that this very woman came out of Mr. and Mrs. Truman? They might have had a child resembling her in many properties. Perhaps in some possible world Mr. and Mrs. Truman even had a child who actually became the Queen of England and was even passed off as the child of other parents. This still would not be a situation in which this very woman whom we call 'Elizabeth II' was the child of Mr. and Mrs. Truman, or so

[12] Kripke, *Naming and Necessity*, 114, fn. 56.
[13] Ibid., 113-114.

> it seems to me. It would be a situation in which there was some other woman who had many of the properties that are in fact true of Elizabeth. Now, one question is, in this possible world, was Elizabeth herself ever born? Let's suppose she wasn't ever born. It would than [sic!] be a situation in which, though Truman and his wife have a child with many of the properties of Elizabeth, Elizabeth herself didn't exist at all.[14]

By analyzing these examples, it becomes obvious that Kripke does not establish a so-called 'principium individuationis', as is suggested in some literature – especially in some German literature – on Kripke.[15] According to this principle, a certain property, if necessary, can only be held by a certain individual. It raises the question whether an individual can be identified (or even transworld-identified) by its essential properties – or more precisely: by its necessary origin. It can be stated without any doubt that Kripke denies a principium individuationis in his analysis of the wooden table example above since he writes in a footnote of Naming and Necessity:

> [I]f the very block of wood from which the table was made had instead been made into a vase, the table never would have existed.[16]

Even though the table has the essential property of being made from the very block of wood, the very block of wood might have been used to produce something different. Being made of the very block of wood does therefore not identify Kripke's table.

Also, the example of Queen Elizabeth II cannot be seen as an example of a principium individuationis. Even though Kripke claims that there is no possible world in which Elizabeth II comes from other parents, her very parents do not make her special among other individuals: Her younger sister Margaret was born of the same parents. Even if we take the more precise wording of "from a totally different sperm and egg" instead of "from different parents"[17], Hägler's

---

[14] Ibid., 112-113.

[15] Cf. Rudolf-Peter Hägler, *Kritik des neuen Essentialismus: Logisch-philosophische Untersuchungen über Identität, Modalität und Referenz* (Paderborn: Schöningh, 1994), and Christof Rapp, *Identität, Persistenz und Substantialität: Untersuchung zum Verhältnis von sortalen Termen und aristotelischer Substanz* (Freiburg and München: Alber, 1995).

[16] Kripke, *Naming and Necessity*, 115, fn. 57.

[17] Cf. ibid., 112-113.

assumption that Kripke established a principium individuationis is simply not true: We might stipulate a possible world in which Elizabeth II has a monozygotic twin sister. Hence it is only a contingent and not a necessary fact that the fusion of the very ovum with the very sperm leads to Elizabeth II. It is solely necessary that Elizabeth II comes from the fusion of the very ovum with the very sperm – or in other words: 'being a result of this very fusion' is a necessary property of Elizabeth II. This is the way we have to understand Kripke's principle of the necessary origin.

However, it fits best to Kripke's intuition-based way of argumentation that he elaborates this principle in a footnote only – and criticizes it himself in the preface, which was added a few years later, as requiring further discussion.[18] As I already adumbrated, it is Kripke's philosophical goal to draw 'better pictures' and not to elaborate fully consistent theories.

The contemporary 'new essentialism' which focuses on the classical Aristotelian substance ontology declines Kripke's necessity of origin. It is claimed that substances (an Aristotelian term which Kripke avoids on purpose[19]) have to be ontologically independent. But how can Kripke's table or Elizabeth II be ontologically independent if individuals come to existence by their necessary origin? This is how we are to understand the following objection by E. J. Lowe:

> I must reject the Kripkean thesis of the 'necessity of origin' – according to which, for example, it is part of the essence of a living organism, such as an elm tree, that it grew from a certain seed. For the seed is presumably not to be identified with the mature tree. However, I do consider this thesis to be mistaken. I am happy to concede, perhaps, that the tree could not have grown from a different seed, but I am not prepared to concede that it did have to grow from this seed, because it seems to me perfectly intelligible to suppose that this very tree could have existed from eternity, or could have come into existence ex nihilo.[20]

But even if we were to accept the necessity of origin as an intuitive principle of essentialism, it does only apply to individuals, but not to natural kinds. However, Kripke also pursues a natural kind essentialism

[18] Cf. ibid., 1.
[19] Cf. ibid., 114-115, fn. 57.
[20] Jonathan Lowe, "Substance and Identity," in *Substanz*, ed. Käthe Trettin (Frankfurt: Klostermann, 2005), 42.

as I already mentioned when I referred to Kripke's famous 'Water is H_2O' example. Kripke derives the natural kind essentialism from his reference theory (if we may call it a theory at all and not a picture!). For Kripke, both proper names and natural kind terms are rigid designators: They refer to the same individual, respectively natural kind, in all possible worlds, given that the very objects exist. Taking this approach to reference theory for granted, individuals do not only have essential properties of their own, but also as part of the natural kind they belong to. For example, a certain drop of water out of the River Thames belongs to the natural kind 'water' – therefore it shares the essential property of being H_2O.

Hereby, it is important to understand that the reference of a natural kind term "depends upon the actual nature of the particular things that serve as paradigms"[21] – as stated by Hilary Putnam who shares similar views as Kripke on reference theory and essentialism. To illustrate the view of both authors by the water example, we can say that the reference of the natural kind term 'water' was fixed by pointing to a certain glass of water (or a certain river, lake, see or whatsoever). Afterwards we find out that either the very water, which fixed the reference, or another paradigm of this natural kind has the property of being H_2O. Since this property is – according to Kripke – essential, we cannot imagine a possible world in which water does not consist of H_2O molecules. Kripke illustrates this by his example of fool's water:

> If there were a substance, even actually, which had a completely different atomic structure from that of water, but resembled water in these respects, would we say that some water wasn't H_2O? I think not. We would say instead that just as there is a fool's gold there could be a fool's water; a substance which, though having the properties by which we originally identified water, would not in fact be water.[22]

But as I see it, the intuitive essentialism of natural kinds is somehow ill-founded: Kripke's essentialism of individuals has at least a more or less plausible principle to find out the essential properties of a certain object – the necessity of origin. But H_2O can hardly be said to be the origin of water – neither is 'being an animal' (to take another of Kripke's

[21] Hilary Putnam, "Meaning and Reference," in *The Journal of Philosophy* 70, no. 19 (1973), 711.
[22] Kripke, *Naming and Necessity*, 128.

examples) the origin of a tiger, even though it is the essential property of every tiger, so to speak something that all tigers have in common. The problem I am targeted on is that Kripke does not give us a single hint how we should find out which of the various properties of a natural kind are essential and which are only contingent. He gives only examples which he believes to be intuitively evident like 'A tiger is necessarily an animal' or 'Water is necessarily H_2O'. But especially the latter example, is in my opinion, problematic: It is not very probable that the man on the Clapham omnibus would regard the statement 'Water is necessarily H_2O' as intuitively evident, not even 'Water is H_2O' without claiming its necessity. For most people, it is probably more intuitive to say that 'Water is liquid' is a necessary truth, particularly because water, as one finds it in nature, is rarely pure H_2O.

However, I am not sure whether it is useful to criticize any particular examples and therefore do not want to continue with that. The main problem of Kripke's essentialism, which I wanted to point out, is that Kripke refuses to name any proper principles to distinguish essential from contingent properties (both for individuals and natural kinds).

Nevertheless, the vagueness of Kripke's intuitive examples was enough to overcome the anti-essentialist paradigm in philosophy – and thus helped to revive Aristotelian essentialism in contemporary discussions. I will come back to this observation in my final outlook.

4. A postcard to Quine: Summary and outlook

In my paper, I tried to develop my point that Kripke revived the Aristotelian idea that individuals and natural kinds have some of their properties necessarily and others contingently. I showed how Kripke rejected the counter-intuitive anti-essentialism which is first and foremost associated with Quine, but has a long tradition going back to Kant[23]. Kripke showed that the anti-essentialists contradict some fundamental intuitions about language use, and he also builds a vague essentialist theory on these intuitions. In the last part of my paper, I

[23] Cf. Richard Rorty "Kripke vs. Kant," in *London Review of Books* 17, no. 2 (1980), 4-5, and Hägler, *Kritik des neuen Essentialismus*.

tried to sketch some general remarks on Kripke's essentialism. It hopefully became clear why its vagueness makes philosophers even today feel as if they were in a metaphysical jungle.

However, even though the 'new essentialism' movement does not totally agree with the views Kripke elaborated in *Naming and Necessity* and his later works, we can well say that Kripke's analysis was a milestone on the way to bring contemporary substance ontology into vogue again. Richard Rorty writes in his review of *Naming and Necessity*:

> Just when it seemed that the [anti-essentialist] dialectic which Kant began had culminated in universal acceptance of the relaxed pragmatism of Wittgenstein and Quine, Kripke exploded his bomb.[24]

In my paper, I tried to analyze the essentialist bomb Kripke exploded by simply following his fundamental intuitions. I am convinced that an impartial and open-minded reception of Aristotle and other important essentialists in history would not be possible today without Kripke – or at least, it would be more difficult to speak about an individual's essential properties. Without Kripke, philosophy today would still be banned to the pragma-empiric jungle of Quine and other early analytics.

Literature

Hägler, Rudolf-Peter. *Kritik des neuen Essentialismus: Logisch-philosophische Untersuchungen über Identität, Modalität und Referenz.* Paderborn: Schöningh, 1994

Hughes, Christopher. *Kripke: Names, Necessity, and Identity.* Oxford: Oxford University Press, 2004.

Kripke, Saul. *Naming and Necessity.* Cambridge: Harvard University Press, 1980.

———. "Identity and Necessity." In *Philosophical Troubles. Collected Papers*, 1-26. Oxford: Oxford University Press, 2012.

Lowe, Jonathan. "Substance and Identity." In *Substanz*, edited by Käthe Trettin, 33-51. Frankfurt: Klostermann, 2005.

Putnam, Hilary. "Meaning and Reference." In *The Journal of Philosophy* 70, no. 19 (1973): 699-711.

Quine, Willard Van Orman. "Three Grades of Modal Involvement." In *The Ways of Paradox and other Essays*, 156-174. New York: Random House, 1966.

[24] Rorty, "Kripke vs. Kant," 4-5.

———. "Two Dogmas of Empiricism." In *From a Logical Point of View: 9 Logico-philosophical Essays*, 20-46. Cambridge: Harvard University Press, 1961.

———. *Word and Object*. Cambridge: MIT Press, 1960.

Rapp, Christof. *Identität, Persistenz und Substantialität: Untersuchung zum Verhältnis von sortalen Termen und aristotelischer Substanz*. Freiburg and München: Alber, 1995.

Rawls, John. *A Theory of Justice*. Cambridge: Harvard University Press, 1971.

Rorty, Richard. "Kripke versus Kant." In *London Review of Books* 17, no. 2 (1980): 4-5.

Stalnaker, Robert. "Possible Worlds Semantics: Philosophical Foundations." In *Saul Kripke*, edited by Alan Berger, 100-115. Cambridge: Cambridge University Press, 2011.

Natural Kind Terms and the Necessary A Posteriori

Adriana Pavić
(University of Göttingen)

In this essay, I am concerned with propositions which we can come to know only via empirical investigation, yet are claimed by Kripke to hold metaphysically necessary, namely propositions expressed by statements referred to by Kripke as "theoretical identities" involving natural kind terms. I tentatively reject anti-realist attacks on the status of these claims as 'necessary a posteriori' on the grounds that the anti-realist fails to do justice to our actual practice of using natural kind terms. In illustrating these points, I appeal to Stroud's reflections on the notion of necessity and Brown's account of the reference of natural kind terms, subsequently amending them by my thesis that the tacit linguistic rules governing the use of natural kind terms involve a priori premises deeply anchored in our conceiving the world which seem to be responsible for the intuitive appeal of the essentialist view put forward by Kripke. Yet these observations do not constitute a positive argument in favor of the purported metaphysical necessity of the relevant propositions. So in the last part of the essay, I rather put up for discussion the question whether upon philosophical reflection on the presuppositions underlying our use of natural kind terms we can soften Kripke's contention concerning the status of propositions like the one expressed by "Water is H_2O" as metaphysically necessary without thereby adopting the anti-realist stance according to which the necessity of such claims is based on a linguistic stipulation. The idea put forward by Putnam (1990) seems to be to regard the relevant proposition's truth as only nomologically necessary on the basis of our inability to assess them for truth in possible worlds in which the most fundamental laws of nature do not obtain.

1. Kripke's contention: empirical access to metaphysical necessity

Through Kripke's Naming and Necessity (N&N), metaphysics "invaded" the philosophy of language, thereby regaining attention among philosophers generally and more specifically entering the realm of analytic philosophy, from which many had wished to exclude it once and for all. Essentialism – the metaphysical doctrine which claims that

objects have essential properties as opposed to accidental ones – came to the fore and with it the notion of metaphysical necessity as opposed to epistemic possibility (conceivability) and semantic necessity (truth by definition, meaning postulate).

According to the picture Kripke advocates, metaphysical claims concerning the status of some statements as necessarily true can be deduced from certain semantic properties of the expressions involved in these statements, such as their rigidity.

The statements he envisages are primarily identity-statements such as "Hesperus is Phosphorus" or "Water is H_2O", though he is also concerned with statements like "cats are animals", whose logical form can be analyzed as a universally quantified conditional.[1]

Now, the discovery that Hesperus and Phosphorus are one and the same heavenly body, just like the discovery that water is composed of H_2O molecules, is the result of empirical investigation. But scientific discoveries like these, according to Kripke, have a modal dimension: If true, they could not have been false, i.e. their truth is metaphysically necessary, so they contradict the traditional assumption that we can gain insight into necessary truths only a priori and they constitute examples of "the necessary a posteriori".

In this essay, I will be concerned with natural kinds and natural kind terms, respectively, not with particular objects such as the Venus, the Queen or a certain table and the corresponding linguistic devices of reference (without assuming that in all the latter cases Kripke's theses hold unproblematically).

Are we indeed committed to the strong metaphysical claims Kripke advocates to account for the evidence he emphasizes? Anti-realists like Canfield come up with an alternative proposal, backed up by a variety of interesting arguments.

In the first section of this essay, I want to point out why, in my opinion, neither of the two rival proposals can yield conclusive arguments against the other one, drawing on Stroud's views on

[1] Kripke subsumes the relevant statement under the potentially misleading header "theoretical identities".

metaphysical necessity presented in his Engagement and Metaphysical Dissatisfaction (2011).

Then, in the second section, I take a look at the semantics of natural kind terms, disclosing a realist picture deeply entrenched in our use of these terms, which partly explains the appeal of Kripke's metaphysics in N&N and casts some doubt on the most radical anti-realistic theses. In this regard the work of Jessica Brown proved helpful. Finally, I propose to weaken Kripke's strong metaphysical claims in favor of a weaker form of realism endorsed by Putnam in his later work.

2. Full-blooded essentialism vs. Wittgensteinian anti-realism?

It seems natural to assume that the substances and kinds we encounter in our natural environment have some properties which are essential to them, so that without having these properties they would not be instances of the respective substance or kind. How else, if not through investigation of these properties, could we be able to distinguish gold from fool's gold, for example?

Natural kind terms, according to Kripke, refer rigidly, i.e. they refer to the same kind in every possible world, and it is up to the scientists to tell us how to determine if a given sample is a sample of a given kind. If scientists succeed in identifying the relevant property, their discovery amounts to an insight in truths which hold in all possible worlds. Kripke writes:

> [S]tatements representing scientific discoveries about what this stuff *is* are not contingent truths but necessary truths in the strictest possible sense. It's not just that it's a scientific law, but of course we can imagine a world in which it would fail. Any world in which we imagine a substance which does not have these properties is a world in which we imagine a substance which is not gold, provided these properties form the basis of what the substance is. (Kripke 1980:125)

It is conceivable that the scientists got it wrong, that it turns out that gold is not Au79. But provided that it is not wrong, it could not have been otherwise. In the same vein Putnam concludes as to water's not

being H_2O: "it is conceivable, but it isn't logically possible!" (Putnam 1975a: 233).[2]

Anti-realists object that Kripke should have said much more about how, according to his view, scientists can discern candidates for an essential property from accidental properties of substances, kinds, and phenomena. Only a certain class of properties seems to come into question, not every property is a candidate for an essential one. Canfield points out that

> Since not all discovered features are essential, a theory of essence is presupposed in passing from the proposition that water is H_2O to the proposition that it is of the essence of water to be H_2O. But what theory? (Canfield 1983:107)

Moreover, for the anti-realist it is not at all so evident that we can discover essential properties in the metaphysically charged sense envisaged by Kripke. They propose to cease using the notion of essence or to deflate it completely. According to the latter option, endorsed by Canfield, scientists just *decide* what is essential, partly due to pragmatic considerations.

The anti-realist has made an interesting proposal, but on what grounds could he persuade us to endorse his view? I shall attend to this question after having clarified the problem he points to.

Essential properties *as such* cannot be discovered by science. Science arguably can discover structural, fundamental, underlying or maybe even *the most fundamental properties* of substances.

The putative *a posteriori* necessity of the relevant statements is a result of combining these findings with certain *a priori* postulates. In the realist picture, the principle that the composition of a substance is essential to it and thus a change in composition is necessarily a change in identity, is a case in point. Once we split the relevant statements into their two components it becomes visible that the *a priori* component imports the necessity, whereas the *a posteriori* component is an ordinary empirical statement. Kripke seems to acknowledge this in emphasizing the conditional nature of our putative epistemic access to the relevant

[2] Putnam does not use the term "metaphysical possibility", but seems to have in mind the same that Kripke has. More on this below.

metaphysical necessities: *if* water is H_2O, *then* it is necessarily H_2O. But he does not give any substantial argumentative support for the relevant *a priori* postulates, in this important respect, he relies solely on intuitions. Now, the implicit *a priori* component amounts to two completely different things in the realistic and in the anti-realist picture; though, as I will try to show, both can be conceived as stipulative in a sense. But "stipulation" amounts to completely different things depending on whether we prefer a metaphysically charged realism or a Wittgensteinian anti-realism.

The claim that the properties in question are essential, and hence are attached to the substance with necessity, is an empirically warranted metaphysical hypothesis in the one case, and a semantic stipulation on the basis of scientific evaluation of empirical evidence in the other case.

Understood in the first (the realistic) sense, the "stipulation" is not easily recognizable as such. To see that we unconsciously do stipulate, we have to reflect on the *a priori* metaphysical premises we bring to bear on the discovered patterns in nature and reveal these premises as unprovable.

The second (the anti-realistic) sense of "stipulation" is an explicitly definitional one: we elevate a certain property, confer it the status of being the essential one. According to the anti-realist picture, this affects the linguistic meaning of natural kind terms, even though the features stipulated as essential are not "analytic", in the sense of being arrived at by conceptual analysis alone.

Accordingly, there are two ways to conceive of the notion of essentiality: the metaphysical one, invoked by Kripke, and the criterial one, proposed by Wittgensteinian anti-realists like Canfield. Two different varieties of necessity come with each picture, metaphysical necessity and semantic necessity, respectively.

But the anti-realist claim that we regard the relevant statements as necessarily true because they are "rules of grammar" in a Wittgensteinian sense, is in tension with our use of the concept of necessity, as Stroud brings out in the following passages:

> [T]o accept such an explanation of necessity we would have to acknowledge that if it had not been for that explanatory factor the necessary truth in question would not have been necessary or perhaps even true. That means we would have

> to grant at least the possibility of circumstances in which something that we regard as necessarily true [···] would not have been necessarily true. But no one could consistently grant such a thing. (Stroud 2011:77)

We do not conceive of necessity as man-made, so to speak; something is necessary if it could not have been otherwise. But rules of language are evidently contingent, as Canfield himself emphasizes when he characterizes his theory as "a form of linguistic relativism" (Canfield 1983:106).

The irreducibility of our notion of necessity, which Stroud emphasizes does not go well together with considering the statements Kripke is concerned with as "rules of grammar".

The *a priori* condition needed for the "theoretical identities", which Kripke considers as examples of the "necessary *a posteriori*" to hold necessarily discloses assumptions deeply entrenched in our thinking about the world. It shows that our judgements aim at nature as it is independent of us, but at the same time raises the skeptical problem that we eventually cannot be sure if we can succeed in representing the substances and kinds in nature as they *really* are.

According to Stroud, as I conceive of his view, this is a case in which neither a positive nor a negative metaphysical verdict is epistemically accessible. There is no proof available, not even in principle, neither for our assumption that substances have an essence, nor for the contrary. But the essentialist preconceptions governing our thinking of and speaking about natural kinds nevertheless seem to be indispensable.

The anti-realist sees himself in a better position, though, for he purports not to rely on any metaphysical speculation. He does not even have to deny explicitly that there is something like essential properties in the metaphysical sense; he can be content with insisting that his opponents' talk of essence simply does not make sense and should be replaced by the criterial sense proposed by Canfield-style anti-realists. But this causes conceptual entanglements due to the irreducibility of the notion of necessity emphasized by Stroud.

However, in the next section I propose to illustrate why only the first picture fits in with how we use natural kind terms. Although the talk of relevant discoveries as discoveries of essence can also be considered as a kind of stipulation by us humans, within the realist framework we do

not *intend* to just stipulate and define, we intend to fit in with facts independent of us.

3. The *characters* of natural kind terms

Kripke does not say much about the linguistic meaning of natural kind terms. He does oes say something about how the reference of these terms is fixed, namely *via* a sort of baptism in which we ostensively define to what kind of entity the relevant term is to be applied. But how is the linguistic meaning of natural kind terms related to the set of all "members" of the natural kind? Let us look at what Putnam has to say about that and contrast his view with the anti-realist view.

According to Putnam, we have to know at least something about tigers to be recognized as one who knows the meaning of the word "tiger". It does not suffice, as he puts it, to know just that tigers are physical objects and nothing more.[3] We have to know some stereotypical feature of tigers, like being cat-like animals with striped fur etc. So far, this is not especially controversial.

None of these stereotypical properties encompassed in the linguistic meaning of the relevant terms constitutes a necessary condition for membership in their extension, none of these features is "analytic", i.e. true of every member of the extension by definition.

But, evidently, there is a link between the linguistic meaning of a natural kind term and its extension, otherwise the respective extension could not be *its* extension. How does Putnam account for this fact?

In the same way as in Kripke's picture, the reference of a natural kind term is fixed ostensively, in the case of water in roughly the following way: "the term 'water' shall refer to all and only instances which stand in the same$_{L(=liquid)}$-relation to this liquid here".[4] (Whether the same$_{L}$-relation obtains, is to be determined by scientists ultimately. It is remarkable that in explaining this mechanism of reference, Putnam makes use of a "weaker" terminology than Kripke. In particular, he talks

[3] Cf. Putnam 1975a:247.
[4] Cf. Putnam 1975a: 225, 231.

about "important physical properties" or "the micro-structure" of a substance instead of its essence.[5]

3.1 Experts shaping meanings?

Striving to find out what a substance "really is", we are obviously not practicing conceptual analysis; we are not engaged in an armchair *a priori* deliberation as, for example, in the case of the word "know" (at least according to most philosophers). In their investigation of substances and kinds, scientists also do not see themselves as enquiring the linguistic meaning of the corresponding terms.

Scientific discoveries can change concepts, but they don't have to. The structural properties, on the basis of which membership in the extension is determined, often do not enter in our common sense concept of the respective kind. Knowledge of these properties is not necessary to understand the linguistic meaning of the relevant natural kind terms.

Putnam insists that non-expert competent speakers have *normal* knowledge, which is completely satisfying and enables one to take part in the linguistic exchange perfectly well. This is underscored by the "Principle of Reasonable Ignorance" he introduces in *Language and Reality*:

> I can know the meaning of the word 'gold' without knowing, explicitly or implicitly, the criteria for being gold (contrary to John Locke), and without having any very clear idea at all just *how* the word is tied to whatever it is tied to. (Putnam 1975b:278)

In contrast, according to the anti-realist view, the discovery that water is H_2O and its stipulation as a criterion for being water generated a change of the meaning of "water". The meaning of a natural kind term changes with a (new) scientific determination of the criterion for membership in the extension of that natural kind term, hence in the anti-realist framework the initial ostensive definition is *not* the basis for reference-fixing. But this view is implausible, for in case the scientists get it

[5] Cf. Putnam 1975a: 232.

completely wrong and later correct themselves, they constantly refer to the same substance and no modification of the meaning seems to have occurred. Other reasons for the implausibility of the anti-realist proposal will emerge once we take a look at Brown's theory.

3.2 Brown's account of the reference of natural kind terms

Jessica Brown provides an expansion, and, in my opinion, an improvement of the Kripke/Putnam view, without giving up the central idea of ostensive reference-fixing. The improvement consists in her clarifying the conditions under which a successful ostensive definition is possible.

The ostensive reference-fixing is based on our general capacity to identify different natural kinds in our natural environment and distinguish them in a fairly reliable way by means of their manifest properties. Presupposing the *recognitional capacity* envisaged by Brown inherently involves a "naïve" realistic element.

According to Brown, there are two requirements to be fulfilled for a person to have a recognitional capacity for a natural kind: (1) She has to have a discriminatory ability with respect to the kind in question, *i.e.* she has to be able to discern that kind from any other kind *in her local environment*. (2) Beyond that she has to appreciate the *metaphysical nature* of natural kinds. We need to import metaphysical assumptions concerning non-manifest properties of natural kinds in order to be able to distinguish cases in which there is only a recognitional capacity for anything with the *appearance* of a certain kind present from cases where a recognitional capacity for that kind itself is present.

This appreciation of the metaphysical nature of natural kinds is to be understood dispositionally, it need not involve a full conceptualization of this nature. The presence of the metaphysical assumption is revealed by our disposition to adjust our judgments in accordance with scientific findings. Brown puts it as follows:

> In order to have a recognitional capacity for silver, a subject needs to have not only a certain discriminatory ability, but also an appreciation of the metaphysical nature of natural kinds. In particular, she needs to appreciate that whether an item is of some particular natural kind is determined by its fundamental

> properties. She needs to appreciate that similarity in the appearance of two items is only defeasible evidence for the claim that there is a natural kind instantiated by both, defeasible in the light of evidence about their fundamental properties. (Brown 1998: 286f.)

The metaphysical assumption Brown points to is also fundamental for our understanding of the linguistic meaning of natural kind terms and is at the bottom of our use of these terms. I want to argue that Brown's condition (2) for having a recognitional capacity for a natural kind forms part of the *character* (in the sense of Kaplan) of every natural kind term and *qua* linguistic rule constitutes a constraint on the understanding of the linguistic meaning of natural kind terms, too. Let me illustrate this further.

In correcting ourselves by taking back a claim to the effect that a given sample is a sample of water, gold or whatever, in the light of scientific evidence to the contrary, we demonstrate our competence in using natural kind terms. Using these expressions, we intend to refer only to items which meet Putnam's "same$_{kind}$"-criterion.

The normative impact of the *character*-rules governing natural kind terms would expose a speaker maintaining "well, all samples of water tested so far consisted of H_2O, but maybe there are others with another underlying structure" as *linguistically* incompetent. It is not just that he or she knows very little about water, it is rather that such a person violates the implicit commitment laid down in our linguistic conventions.

This general commitment is already in force because the scientists are still in the dark with respect to the underlying features of a given substance. Let us assume that we have identified a particular sort of metal, *i.e.* we are able to discern samples of this metal as samples of the same substance *via* some appropriate "operational definition". We have ostensively introduced the term *m* for this substance, because science has not yet investigated the underlying structure of the substance. At this particular time already (when we are yet in search of a candidate for an essential property) we consider the following sentence as false under every possible substitution for *F*, on the assumption that *F is* an essential property:

> (S) Most, but not all samples of *m* are *F*.

If F is an essential property, then any sample which lacks that property cannot be a sample of *m*.

Realizing this is gaining insight in the semantics of natural kind terms. As competent speakers we need not know what the fundamental property of a given kind is. Nor do we even have to be aware of the metaphysical commitments underlying our use of natural kind terms, just as little as we have to be aware of the rules forming part of the *character* of other linguistic expressions we use every day, for example the rule that "I" refers to the speaker. This is in accordance with what Putnam claims in his "Principle of Reasonable Ignorance" quoted above.

Let me summarize what I have said about the relation between the meaning and the extension of natural kind terms. We can draw a distinction between an external and an internal component in the meaning of natural kind terms, *the external* one being primarily responsible for reference-fixing, in that it demands an ostensive reference to paradigmatic samples, which in turn makes an exact determination of the extension in the future possible.[6]

Superficial properties also play a role in fixing the reference, for they are the basis for the pre-scientific individuation of kinds by means of our recognitional capacities. The relevant manifest properties enter in the conventional meaning of the term – they are precisely the features which constitute the stereotype in Putnam's sense. So Brown's first condition for having a recognitional capacity for a natural kind and Putnam's stereotype can also be correlated.

The internal component involving the a priori assumptions also plays a role in reference-fixing, albeit an implicit one: The *character*-rules secure the usefulness of the ostensive definition (demanded by the external component) by way of ascertaining that the extension be determined by reference to underlying, structural properties of the paradigmatic samples, so, in a sense, it can be regarded as the "linguistic trigger" for searching the essential property of a kind.

---

[6] This does not mean that we cannot introduce a natural kind term descriptively, but a descriptive reference-fixing like "the substance which caused x" would be regarded as preliminary, for in such cases we are in anticipation of becoming acquainted with this substance *via* perception, being able to examine it scientifically etc. in the future.

Consequently, semantic externalism should be retained; the reference of natural kind terms depends "on the world", not on mental entities of any kind. It depends on the world in a way which is determined by the ostensive definition of the term and the prescriptions provided by the *character* of natural terms in general.

Our thinking of, and dealing with, natural kinds, substances, and phenomena stands on a metaphysical footing which leads us to consider scientific discoveries of underlying properties as discoveries of essence. As a result, we are willing to accept the modal consequences of such discoveries emphasized by Kripke.

So in my view the scientific use of natural kind terms is not disconnected with our everyday use (contra Putnam 1990, but in accordance with Putnam 1975a, where he emphasizes the role of "experts" in the linguistic community). This allows us to bridge scientific curiosity and the pragmatic requirements of daily life in a unifying picture of our conception of natural kinds.

4. Putnam's revision: Does nomological necessity (in a realistic framework) suffice?

After this sketch, let us take a look at an example which threatens to disarrange our previous conclusions.

According to our scientific knowledge of the world, no substance can change when left on its own.

There has to be an impact from outside, so to speak – an extraneous cause bringing about the change in the substance, for example its passing from the solid state of aggregation into the liquid one through heating.[7] Another example is chemical reactions between substances: Substances merge and their chemical composition changes, whereby new substances emerge. At any rate, every change concerning a substance has to have a cause, has to be caused by something.

[7] Of course, we can, and do, say that the capacity or disposition to undergo certain changes pertains to the substance itself; it is not just that the cause (for example heat) changes the substance, we can just as well say that the substance changes due to certain circumstances/influences/causes. The point remains that no substance changes "out of the blue".

But nomological impossibility does not imply metaphysical impossibility. There could have been other fundamental physical laws. So the following scenario should be classified as metaphysically possible.

The Strange Mutation

Imagine a big lake in Finland filled with crystal clear blue water – a paradigmatic instance of the kind. If you feel like it, you can even assume that the liquid in this lake was one of the samples by reference to which water was ostensively defined a long time ago. But now something very strange occurs: the liquid in the lake begins to alter chemically. *The content of this very lake* undergoes a change in chemical composition resulting in the molecular structure XYZ, although no observable feature of the environment has changed.[8] During this transformation, which you can imagine as occurring very slowly or instantly, all the perceptually observable qualities of the substance remain the same.

What should we say in this case? It is difficult. We are not prepared to deal with such cases in our language use, because such cases are nomologically impossible, they contradict the most basic laws of nature. The strange occurrence seems to be conceivable, otherwise you could not understand the example. And I see no argument against its being metaphysically possible either. There could have been other laws of nature allowing changes to occur without a cause, a world without such things as "fixed substances", *i.e.* substances with an underlying structure constant over time. That we humans would have a hard time getting along in such a world (to put it mildly) does not make it metaphysically impossible. So, if it is metaphysically possible, should we say that in this scenario water changes its composition or rather that water transforms into something else which is no longer water?

This example is based on the metaphysical assumption that for every natural kind there is a distinctive feature which constitutes its

[8] It is important to keep in mind that in this example we stipulate that there is no cause for this change of molecular structure, it is not just that we have not detected it yet (for example, an extraterrestrial gas which has entered our atmosphere unnoticed or something like this).

essence to its limits. Is the statement "water is H_2O" truth-evaluable in a world in which the fundamental laws of nature are not in force?

We would have no use for this sentence with respect to such a world, for its meaning rests upon the possibility of classification of substances. Whether we regard them as metaphysically necessary or not, our criteria for distinguishing substances are applicable only to worlds in which the present laws of nature are in force; we do not possess distinguishing criteria for substances independently of this condition. The relevant statements express necessary truths against the background of a world in which substances have a constant underlying structure whose integrity can be affected by causal means only. But other worlds are conceivable (at least in a sense) and nothing speaks against their being metaphysically possible in principle.[9]

Let w_n be the set of all nomologically possible worlds. Should we say that it is metaphysically necessary that water is H_2O *relative to* w_n? Can something be metaphysically necessary only relative to worlds in which our laws of nature obtain? This cannot work, for it is just the gist of the idea of metaphysical necessity that it holds independently of all contingent factors, even the most fundamental ones like the laws of nature. So we are not allowed to restrict the domain of worlds to those which are nomologically possible.

Are the criteria for membership in the extension of a natural kind term merely nomologically necessary then? Would water's not being H_2O be incompatible with the fundamental laws of physics? This question seems to be on the wrong track, for to answer it, we already need to have available a criterion for determining whether a given sample is a sample of *water*. And precisely this criterion is "being composed of H_2O molecules"; we would not call anything water unless it were composed of H_2O,[10] so the question whether this would be compatible with our laws of nature does not arise:

---

[9] Of course, one could argue that the laws of nature themselves are not contingent, *i.e.* that there could *not* have been other ones, but this seems to muddle up the concept of metaphysical modality. I cannot pursue this issue here, but Sidelle (2002) convincingly argues against this proposal.

[10] I ignore problems of vagueness resulting from the fact that normally we do not encounter chemically pure water in our natural environment, for they do not have a bearing on the point presently under discussion.

> Once we have discovered that water (in the actual world) is H_2O, *nothing counts as a possible world in which water isn't H_2O.* (Putnam1975a: 233)

This seems to be preserved in spite of the mentioned difficulties, for even a world in which completely other laws of nature are in force is *prima facie* no world in which water is not H_2O, though it might be a world in relation to which talking about water does not make sense because we do not have criteria for the individuation of substances applicable in such a world, or simply a world in which there is no water. So the following possible reactions are available in cases like *The Strange Mutation* example:

(1) The question whether there is water in such a world does not arise.

(2) There is no water in such a world.

(3) There is water in such a world when and only when H_2O molecules are instantiated (regardless of what happens before or afterwards).

(3) seems to be the Kripkean response, but it gives the impression of an obstinate application of natural kind terms, whereas (1) and (2) resist applying them in nomologically impossible worlds, thereby suggesting that their application presupposes that some background conditions obtain, namely the continuity of substances in the absence of causal impact etc. This would explain our difficulty in yielding clear-cut intuitive responses to such scenarios.

Thus, maybe on reflection, we can weaken the assumption underlying our understanding and use of natural kind terms to the effect that what we, in fact, have to assume is that there are fundamental properties of every substance conceived as properties which apply in every possible world *in which the same or very similar laws of physics obtain, i.e.* every *nomologically* possible world.

This is, in effect, Putnam's later stance in his paper *Is Water Necessarily H_2O?* (1990). There he qualifies some of his views in *The Meaning of 'Meaning'* (1975). As a result of thinking through some of the consequences of Kripke's claims, he departs from Kripke at some points where this was not apparent in his previous work. This new way promises to be more tangible and still compatible with our use of

natural kind terms and our intuitive judgements of various counterfactual scenarios. So Putnam concludes:

> I now think that the question, "What is the necessary and sufficient condition for being water *in all possible worlds*?" makes no sense at all. And this means that I now reject "metaphysical necessity". (Putnam 1990:70)

This presents itself as a reasonable conclusion in the light of my *Strange Mutation* example and other examples provided in Putnam 1990. Such examples suggest that everything over and above nomological possibility (accessible only *a posteriori*) and logical possibility (accessible only *a priori*) creates confusion, and we can do without metaphysical necessity.

But Kripke could counter along the following lines: "In any world in which there is no H_2O, there is no water either, and even in a world in which substances appear and disappear out of the blue, while H_2O molecules are present, water is present, no matter what happens in advance. So we have a criterion valid even in nomologically impossible worlds!"

Can we stay in accordance with the rules implicit in the linguistic meaning of natural kind terms and still succeed in departing from Kripke's essentialism with its strong metaphysical commitments?

At this point, I want to leave it open for further reflection whether "necessity in the strictest possible sense" in the case of the pertinent statements concerning natural kind is just a nomological necessity under the condition that we have identified the most fundamental properties of the relevant substance. Then we would have: For any property *F* and natural kind *m*: if *F* is a fundamental property of *m*, it is nomologically necessary that *m* is *F*. But for the reasons mentioned above, I must confess that I am not quite sure about that.

5. Conclusion

Canfield is right in emphasizing that in the interpretation of the unusual scenarios discussed by Kripke and Putnam, we go beyond our everyday use of natural kind terms. Nonetheless, we have to explain why, even among the different available choices which Canfield points to in his examples, the majority of speakers would intuitively choose the ones in

accordance with the picture on which Kripke and Putnam rely. I pointed to a possible explanation appealing to the *characters* of natural kind terms, which has to be spelled out in more detail elsewhere.[11]

To get to the theoretical identities' obtaining, which is metaphysically necessary, an *a priori* premise or condition is needed. Kripke presupposes it, without making it explicit or motivating it, apart from his repeated appeals to the intuitiveness of the resulting judgements. These intuitions come with our understanding of how natural kind terms function; they indicate the implicit knowledge of the semantic rules which competent speakers possess. Therefore Canfield's criticism that "he attempts to turn plain facts about kind terms into a philosophical theory of them" (Canfield 1983: 109) is quite comprehensible, in spite of Kripke's insisting that he merely wants to provide a picture, not a theory.

Nonetheless, the common perception of Kripke's work, to the effect that his philosophy of language – particularly his thesis that proper names and natural kind terms are 'rigid designators' – has wide-ranging epistemological and metaphysical implications, is accurate in an important sense.

The background, indispensable in our conceptualizing the world, is reflected in our language. Linguistic conventions comprise many tacit rules, some of which, as I tried to show, express metaphysical assumptions which we maintain *qua* human beings.[12]

Literature

Brown, Jessica. "Natural Kind Terms and Recognitional Capacities", *Mind* 107 (1998): 275-303.

[11] For reasons of space, this is not possible in this paper and will be done in a forthcoming one, where I will try to assess the idea that natural kind terms are indexical expressions introduced by Putnam 1975a, criticized by Burge 1982, and reinforced, for example, by Haas-Spohn 1997. Furthermore, I intend to deal with two-dimensional accounts of the semantics of natural kind terms put forward by Jackson and Chalmers and criticized, for instance, by Nimtz, who (to my mind justifiably) insists that "natural-kind terms do not have senses that give rise to substantial *a priori* knowledge" (Nimtz 2004:145).

[12] In developing many lines of thought I started out from Chapter 6 in Lange 2010. Thanks for the fruitful discussions!

Canfield, John. "Discovering Essence." In: *Knowledge and Mind. Philosophical Essays*, edited by Carl Ginet & Sydney Shoemaker. Oxford: Oxford University Press, 1983.

Haas-Spohn, Ulrike. "The Context Dependency of Natural Kind Terms." In: *Direct Reference, Indexicality and Propositional Attitudes*, edited by Wolfgang Künne, Albert Newen, and Martin Anduschus, 333-349. Stanford: CSLI Publications, 1997.

Kripke, Saul. *Naming and Necessity*, Oxford: Basil Blackwell, 1980 (first published in: *Semantics of Natural Language,* edited by G. Harman and D. Davidson, D. Reidel Publishing, 1972).

Lange, René. *Starrheit, Notwendigkeit und Natürliche-Art-Ausdrücke in Kripkes "Naming and Necessity"*, Magister-Thesis, University of Göttingen, 2010.

Nimtz, Christian. "Two-Dimensionalism and Natural Kind Terms", *Synthese* 138 (2004): 125-148.

Putnam, Hilary. "The Meaning of 'Meaning'." In: *Mind, Language and Reality. Philosophical Papers, Volume 2*, 215-271, Cambridge: Cambridge University Press, 1975a.

———. "Language and Reality", in: *Mind, Language and Reality. Philosophical Papers, Volume 2*, 272-290, Cambridge: Cambridge University Press, 1975b.

———. "Is Water Necessarily H_2O?" In: *Realism with a Human Face*, edited by James Conant, 54-79, Cambridge: Harvard University Press, 1990.

Sidelle, Alan. "On the Metaphysical Contingency of Laws of Nature." In: *Conceivability and Possibility*, edited by Tamar Szabó Gendler & John Hawthorne, 308-336, Oxford: Oxford University Press.

Stroud, Barry. *Engagement and Metaphysical Dissatisfaction. Modality and Value*, Oxford: Oxford University Press, 2011.

Consistency as a tool for interpreting the Philosophical Investigations

Leon Weber
(University of Bamberg)

Interpreting Wittgenstein's work is not an easy task. This becomes apparent when considering the debate on how the *Tractatus logico-philosophicus* is to be understood. In addition to the *Tractatus*, the *Philosophical Investigations* also do not suggest an indisputable interpretation. Therefore, in this paper, I will present an argument for Diamond's reading of the *Investigations*. I will argue that a certain set of readings implies a self-contradiction, while Diamond's approach succeeds in avoiding this: Kripke's understanding of the Investigations will serve as an example for those self-contradictory readings.

1. Kripke's reading

Let me first depict Kripke's exegesis of the *Investigations*. For this, I will follow Stegmüller's account. Kripke uses the rule-following paradox as a key to the *Investigations*. This paradox can be paraphrased by three questions:

> (I) What constitutes the fact that I am following *this rule* rather than another one? [...] (II) What constitutes the fact that *I* am referring to something specific by using some word like 'green' or 'pain'? (III) What constitutes the fact that I have understood some term like 'green' or 'pain'?[1]

According to Kripke, Wittgenstein asks which fact constitutes rule-following. This can be understood as a demand for a set of necessary and sufficient conditions. If and only if those conditions are satisfied, could one be justified in calling something 'rule-following'. Kripke's position is that there is no such finite set of conditions.[2]

[1] Wolfgang Stegmüller, *Kripkes Deutung der Spätphilosophie Wittgensteins: Kommentarversuch über einen versuchten Kommentar* (Stuttgart: A. Kröner, 1986), 10. My translation.
[2] Cf. ibid.

In his view, Wittgenstein denies the existence of such a set, because the way in which we learn rule-following does not allow us to conceive of all of its necessary and sufficient conditions. He gives the rule of addition as an example:

> Let me suppose, for example, that '68 + 57' is a computation that I have never performed before. Since I have performed – even silently to myself, let alone in my publicly observable behavior – only finitely many computations in the past, such an example surely exists. In fact, the same finitude guarantees that there is an example exceeding, in both its arguments, all previous computations. I shall assume in what follows that '68 + 57' serves for this purpose as well. [...] Now suppose I encounter a bizarre sceptic. [...] After all, he says, if I am now so confident that, as I used the symbol '+', my intention was that '68 + 57' should turn out to denote 125, this cannot be because I explicitly gave myself instructions that 125 is the result of performing the addition in this particular instance. By hypothesis, I did no such thing. But of course the idea is that, in this new instance, I should apply the very same function or rule that I applied so many times in the past. But who is to say what function this was? In the past I gave myself only a finite number of examples instantiating this function. All, we have supposed, involved numbers smaller than 57.[3]

In everyday life, we learn the rule of addition by considering single instances of its application. The set comprised of all those instances can never contain every possible addition. There are always instances of rule-following which have not been actualised yet. Considering addition, this means that there is always a pair of numbers which has never been added by anyone and thus, no one could have learned how to add those specific figures. It is obvious that this paradox is not restricted to the domain of mathematics, but extends to every context in which the notion of rule-following is important.

This paradox has serious consequences for our understanding of language. One of our most fundamental assumptions concerning verbal communication seems to be that a given word stands for the same entity in all – or at least, in most – situations. If a speaker A uses a word x at one point in time t_0, then the assumption would be that the meaning of x stays the same in later points of time t_y. But since he learned how to use x in a finite set of situations and the context can be dramatically altered at t_y, the use of x is an instance of the rule-following paradox. There is no way to determine if A follows a certain rule by using x and

[3] Saul Kripke, *Wittgenstein on rules and private language: an elementary exposition* (Oxford: Blackwell, 1982), 8.

thus, communication appears to be impossible. This seems to imply that there is no other option but to accept a radical nihilism when it comes to concepts and meanings.[4]
However, Kripke's Wittgenstein shows us an alleged way out of this nihilism:

> The entire 'game' [...] – that the community attributes a concept to an individual so long as he exhibits sufficient conformity, under test circumstances, to the behavior of the community – would lose its point outside a community that generally agrees in its practices. If one person when asked to compute '68 + 57' answered '125', another '5', and another '13', if there was no general agreement in the community responses, the game of attributing concepts to individuals – as we have described it – could not exist.[5]

In Kripke's opinion, the private language argument gives us reason to think that language cannot function at all if there is not a community of speakers that defines the practices in which the use of language makes sense.[6] Since we obviously have access to language, there must be a community that gives us the ability to engage in linguistic practices and consequently, to follow rules. The truth conditions are replaced by conditions of warranted assertion, grounded in the circumstances under which people do, in fact, attribute sensations to themselves and to others.[7] So, the private language argument helps to conclude that the presence of a community allows speakers to use language despite the rule-following paradox.

3. Self-contradiction

However, this reading leads to a self-contradiction in the *Investigations.* Wittgenstein states:

> Philosophy simply puts everything before us, and neither explains nor deduces anything.--Since everything lies open to view there is nothing to explain. For what is hidden, for example, is of no interest to us. One might also give the

[4] Stegmüller, *Kripkes Deutung der Spätphilosophie Wittgensteins*, 80.
[5] Kripke, *Wittgenstein on rules and private language*, 96.
[6] Cf. Derek McDougall, "The Role Of Philosophical Investigations § 258: What Is 'The Private Language Argument'?," in *Analytic Philosophy*, no. 1 (2013), 71.
[7] Cf. ibid., 66.

> name 'philosophy' to what is possible before all new discoveries and inventions.[8]

Here, Wittgenstein explicitly states that philosophy does not explain or conclude anything. Obviously, for him, philosophy is a purely descriptive enterprise. Consequently, philosophizing does not serve a critical function. If this statement is considered in conjunction with Kripke's reading of the rule-paradox and its sceptical solution via the private language argument, then it becomes apparent that, for Kripke, the different language games in the community of speakers cannot be hierarchical. There cannot be a set of language games that is more justified than another one, because otherwise the notion of 'more justified' language games would serve a critical function – inasmuch as they would serve to criticize the 'less justified' ones – and this is explicitly excluded by the passage from the *Investigations* quoted above. Hence, following Kripke's reading there can be no privileged set of language games. Hösle identifies this assumption as a self-contradiction:

> How can someone be justified in stating that every language game works well without any philosophical regulation while fiercely criticizing traditional philosophy? And how can one state that 'philosophy leaves everything as it is' without self-contradiction, while this statement is obviously directed against those language games that aim to alter things in philosophy?[9]

The statement that there is no set of privileged language games can itself be understood as a criticism of all those language games which privilege some language games. Thus, all of the purely descriptive 'conservative' language games which do not privilege other language games can be considered to be privileged. Since the existence of a set of privileged language games was denied in the initial statement, it is self-contradictory. And because this initial statement follows from Kripke's reading of the *Investigations*, the interpretation leads to a self-contradiction.

---

[8] Ludwig Wittgenstein, *Philosophical Investigations*, translated by G. E. M. Anscombe (New York: Macmillan, 1963), §126.

[9] Vittorio Hösle, *Die Krise der Gegenwart und die Verantwortung der Philosophie* (München: C.H. Beck, 1997), 84. My translation.

3. Diamond's Reading

Still, there are readings of the *Investigations* that do not lead to aforementioned self-contradiction. One of those is Diamond's exegesis, as I will argue. According to Hunziker, Diamond opposes Kripke's reading:

> The ramified debate concerning Wittgenstein's concept of rule-following illustrates the tendency of the metaphysical spirit to misread such statements made by Wittgenstein. They are conceived as being transcendental philosophy and thus claiming that the capacity for language [...] is a necessary condition for things like hope. Kripke's conception of the 'rule-following paradox' [...] is a particularly prominent example for this phenomenon: If one could describe the meaning of 'meaning' by referring to assertibility conditions, then it would be possible to give a theory of meaning that would be rather independent of the 'whirl of life'. It would be independent of the form of life in which the word or sentence is uttered.[10]

She rejects the understanding of the *Investigations* as a work of transcendental philosophy in which the conditions of the possibility of language are laid down. Rather, she reads it as a comprehensive criticism of metaphysics. For her, a central feature of metaphysics is the notion 'must'.[11] She thinks of metaphysics as the form of inquiry which examines entities by laying down necessary conditions for the possibility of the thing. In her opinion, the *Investigations* oppose such a way of philosophizing and the text seems to support that hypothesis:

> For we can avoid ineptness or emptiness in our assertions only by presenting the model as what it is, as an object of comparison—as, so to speak, a measuring-rod; not as a preconceived idea to which reality must correspond. (The dogmatism into which we fall so easily in doing philosophy.)[12]

Here, Wittgenstein explicitly states that a certain emptiness of assertions can be avoided when one does not include necessary conditions in it. For Diamond's Wittgenstein, the alternative to this metaphysical way of doing philosophy is attentiveness towards the role which the subject of

---

[10] Andreas Hunziker, *Das Wagnis des Gewöhnlichen: Ein Versuch über den Glauben im Gespräch mit Ludwig Wittgenstein und Stanley Cavell* (Tübingen: Mohr Siebeck, 2008), 30-31. My translation.
[11] Cf. ibid., 12-13.
[12] Wittgenstein, *Philosophical Investigations*, §131.

inquiry plays in our forms of life. She elucidates this by giving the example of hope:

> [W]hen he [Wittgenstein; L.W.] says that only those who have mastered the use of a language can hope [...], he does not mean that hoping is a logical achievement of some sort, dependent upon mastery of language, but rather that, as he puts it, the phenomena of hope are modes of this complicated form of life' (that is, of life with language)[13].

The mentioned passage by Wittgenstein can be read in at least two ways. Either one assumes that hoping is only possible for those beings which have mastered the use of language or he simply wants to draw attention to the fact that hope plays a role in our forms of life as talking beings. Kripke's interpretation would probably suggest the former understanding, because he inquires into language (or rule-following) by exhibiting necessary conditions for its existence. Diamond opts for the alternative and reads the remark as a nudge to consider the role that hope plays in our forms of life. She suggests that this way the desire for further analysis of the subject of inquiry will fade.[14] Reconsidering the example of hope, this would mean that if we realize which role hope plays in our forms of life, then we would be contented with the insights gained, and the motivational basis of further inquiries using transcendental philosophy would have vanished.

4. No self-contradiction

This should be sufficient to explain how Diamond's reading of the *Investigations* avoids the inconsistency identified by Hösle. For her, the *Investigations* simply are not about non-privileged language games. They are concerned with the understanding of certain concepts by looking at the forms of life in which those concepts play a role. This attentiveness towards forms of life makes further inquiries obsolete. This explains how the *Investigations* can have an obviously critical function and still claim that philosophy leaves everything as it was:[15] Wittgenstein observed that a certain approach towards philosophical

[13] Diamond, "Wittgenstein and Metaphysics," 21.
[14] Cf. Hunziger, *Das Wagnis des Gewöhnlichen*, 11.
[15] Cf. Hösle, *Die Krise der Gegenwart und die Verantwortung der Philosophie*.

topics (the analysis of forms of life) makes our desire for further philosophical investigation go away. In this way, it is critical of other types of philosophy by leaving everything as it was. On this reading, the self-contradiction has vanished.

Measured against consistency, Diamond's reading of the *Investigation* is superior to Kripke's. She avoids the inconsistency which Kripke's exegesis necessarily implies. Maybe there are other considerations that would favor Kripke's approach. This paper is not supposed to settle the question of which is the right interpretation of the *Investigations*. It is supposed to give an argument for Diamond's reading. As already mentioned in the footnote above, I am not sure if the contrast between Diamond's and Kripke's reading is as sharp as Diamond thinks. Some passages in Kripke's *Wittgenstein on rules* seem to suggest that his interpretation can be understood as a therapeutic approach as well. Still, I think that Hösles argument poses a challenge for every reading of the *Investigations* that is metaphysical in the sense mentioned above.

Literature

Diamond, Cora. "Wittgenstein and Metaphysics," in *The Realistic Spirit. Wittgenstein, philosophy and the mind.* 13-37. Cambridge, Mass: MIT Press, 1991.

Hösle, Vittorio. *Die Krise der Gegenwart und die Verantwortung der Philosophie.* München: C.H. Beck, 1997.

Hunziker, Andreas. *Das Wagnis des Gewöhnlichen: Ein Versuch über den Glauben im Gespräch mit Ludwig Wittgenstein und Stanley Cavell.* Tübingen: Mohr Siebeck, 2008.

Kripke, Saul A. *Wittgenstein on rules and private language: an elementary exposition.* Oxford: Blackwell, 1982.

McDougall, Derek A. "The Role Of *Philosophical Investigations* § 258: What Is 'The Private Language Argument'?" *Analytic Philosophy*, no. 1 (2013): 44-71.

Read, Rupert J., Matthew A. Lavery: *Beyond The Tractatus Wars.* London: Routledge, 2010.

Stegmüller, Wolfgang. *Kripkes Deutung der Spätphilosophie Wittgensteins. Kommentarversuch über einen versuchten Kommentar.* Stuttgart: A. Kröner, 1986.

Wittgenstein, Ludwig. *Tractatus logico-philosophicus.* Frankfurt am Main: Suhrkamp, 1963.

———. *Philosophical Investigations.* Translated by G. E. M. Anscombe. 3rd ed., New York: Macmillan, 1968.

Metaphysics for limited beings: how scientific practice leads to methodological concerns in metaphysics

Mads Goddiksen
(Aarhus University)

In this essay, I discuss a particular approach to philosophy that is sometimes labeled metaphysics. The aim of this approach is to unite all human knowledge into one grand logical system which is constructed through a search for the most general metaphysical truths and their relation to all other truths. I will argue that this approach has severe weaknesses. The result that we could hope to gain from this approach is a better understanding of the internal relations in scientific knowledge and the relation between scientific knowledge and other types of knowledge. However, I will argue that the methodology that this approach employs prevents any deep understanding of scientific knowledge. Fitting scientific theories and laws into logically consistent systems of propositions can only be done at the cost of severe distortion and loss of understanding. The final part of the essay therefore suggests a different approach to metaphysics that is more suitable for limited beings

1. Introduction

Every field has its favorite examples. They appear repeatedly in papers and textbooks and are passed on from teacher to student as illustrations and partial definitions of concepts and problems. Philosophy is no different. In the philosophy of mind, the examples tend to be bloody and action packed. If I try to punch my enemy in the nose, but miss, will it then be an intentional act? If I do not miss, will my enemy's feeling of pain be equivalent to, supervenient on, or quite different from a brain state?

In other fields within philosophy, the examples are a bit duller. Those who work with truth theory seem to have an obsession with cats (or is it mats?), and in metaphysics there seems to be a particular preference for the following sentence:

Water is H_2O.

There is much that can be, and has been, discussed about this sentence. Is it a necessary truth? And, if so, what kind of necessary truth?[1,2] Is it a metaphysical truth? Or is it based on an even more fundamental truth about "the very nature" of water? Or even the very nature of being?[3]

I will not venture into the details of these discussions. Instead, I will challenge the basic assumptions of these discussions by pointing to some fairly obvious lessons about general statements like "Water is H_2O", which can be drawn from everyday experience. These will be supplemented with lessons from the philosophy of science about scientific theories and human understanding. In Section 4, I will discuss the implications these lessons have for current discussions about what the aims and methods of metaphysics should be.

2. Stating the obvious

The questions raised by metaphysicians about the sentence "Water is H_2O" mainly concern *what kind of truth it states.* These discussions seem to miss a very important fact:

The sentence "Water is H_2O" does not state *a truth* at all![4]

The sentence can be read in at least two different ways. Firstly, we can read it as a statement of *identity:* "Water is (the same as) H_2O." This is the way it is used in most contexts and certainly the way Putnam reads it when he discusses whether or not water could be something other than H_2O on Twin Earth.[5] We can also read it as stating a *property* of water. This would make the sentence analogous to the sentence "snow is white". While this reading might make the sentence true, it will also fail

[1]Hilary Putnam, "Meaning and Reference," *The Journal of Philosophy* 70, no. 19, Seventieth Annual Meeting of the American Philosophical Association Eastern Division (Nov. 8, 1973), 699-711.
[2]Saul A. Kripke, *Naming and Necessity*, reprinted ed. (Oxford: Blackwell, 1988).
[3]Kit Fine, "What is Metaphysics?," in *Contemporary Aristotelian Metaphysics*, ed. Tuomas E. Tahko (Cambridge: Cambridge University Press, 2012), 8-25.
[4]Hasok Chang, *Is Water H_2O?: Evidence, Realism and Pluralism* (New York: Springer, 2012) reaches a similar conclusion based on an illuminating historical and philosophical discussion.
[5]Putnam, *Meaning and Reference*, 699-711.

to represent the use of the relation between "water" and "H_2O". In science classes, for instance, the teacher might say "water", but (s)he writes "H_2O" on the blackboard.

Reading "water is H_2O" in the usual way, i.e. as a statement of identity, makes the sentence *false* in a world governed by a two-valued logic - i.e. a world where a statement that is not absolutely true is false. This should be obvious to any student in metaphysics who ventures down into the messy world of the senses even for a brief moment.

As an illustration we can consider the following piece of fiction:

Not so long ago there was a very famous professor in metaphysics who traveled the world and gave lectures on the nature of all being. As an experienced lecturer, he knew that it was important to have concrete things and examples to point to while presenting his clear cut, precise and very abstract arguments. Therefore, he always carried an empty bottle with him on his travels (airport security is so tight these days!). Just before giving his lecture, he would find a tap and fill the bottle with water. When he got to the point in his lecture where an example was needed he could point to his bottle and say, "In this bottle I have some water, and, as we all know, water is H_2O." No one would usually object to this.

Of course, if the famous professor had opened the bottle and tasted the liquid inside it, he could easily have convinced himself that water is not *just* H_2O. In some places, he would be able to taste that the liquid contained not only hydrogen and oxygen, but also chlorine. He might even be able to taste that the amount of calcium in the water differs from one region to another (and if not, to see it, he need only compare the showerheads in a couple of the different hotel rooms he stayed in). The downside to this empirical adventure is, of course, that it might be dangerous; in some places the tap water is safe to drink, while in other places it can give you a very painful reminder of the material aspects of the human body.

3. What is Water ?

This false tale reminds us that if we take samples of water from different places and submit them to thorough chemical analysis, the results will not be identical. Certainly, most of the samples will contain mainly oxygen and hydrogen, and most will contain twice as much hydrogen as oxygen. The majority of these atoms will also most likely be bound together in H_2O molecules, but that does not change the fact that the sample will almost certainly contain other things than H_2O. Thus, the statement that "water is H_2O" is not an unconditional truth at all, but a convenient approximation that is useful in most contexts.[6] An important thing to learn when learning to use such approximations is when *not* to use them, i.e. know when the approximation *fails*. H_2O has some very beneficial properties: it can quench your thirst, clean your equipment, even the more delicate parts of it, and it is highly useful as a solvent, for instance in a chemistry lab. Liquids that can be referred to as water both in daily and scientific discourse *usually* have the same properties. However, if you are lost at sea, you should take care to note that the water that surrounds you will only make you *thirstier*. Furthermore, anyone who has spent time in a chemistry lab knows that using the wrong kind of water may ruin the *experiment* or even damage the *expensive equipment*. It is, thus, not very difficult to find contexts where it would be wrong, even dangerous, to treat water as identical to H_2O.

3.1. Objections

The famous professor in metaphysics, and others, might object to all this nonsense. Surely water is H_2O, and what the professor had in his bottle was just water *with something in it*.

This objection rests on an unfaithful reconstruction of what happened when the professor filled his bottle. When he turned on the tap and filled up his bottle, he filled it up with water, not water with

[6]Wimsatt would call it a heuristic. William C. Wimsatt, *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality* (Cambridge, MA: Harvard University Press, 2007). Cartwright would say that the law is true only *ceteris paribus.* Nancy Cartwright, *How the Laws of Physics Lie*, repr. ed. (Oxford: Clarendon, 1984).

added substances. That is what anyone, even a chemist, would say in the given context. To say that he filled it with water with added substances would be to distort the ordinary use of the word 'water' and would force us to draw absurd conclusions. For instance, it would then be wrong to say that humans drink water, since water (i.e. H_2O *without* added substances) is not found anywhere on the planet, except perhaps in laboratories.

So what the professor filled into his bottle was water. However, I will grant that if the bottle with the same content were put in a different context, for instance in the hands of a chemist in her lab, it *would* contain water with added substances. This is because the adequate description of the content of the bottle is not *solely* determined by the content of the bottle; it also depends on the context in which the bottle and its contents are *used*.[7,8] As mentioned before, the heuristic that "Water is H_2O" is safe to use in most contexts, but in the context of a chemistry lab it is not safe. Here a slightly different heuristic may apply: d*emineralized* water is H_2O. But, of course, even this heuristic can fail, and the chemist needs to learn when and how to act if it does.

So it turns out that what metaphysicians and others have treated as a simple, even necessary, truth is not a general truth after all. The problem here is not the choice of example! Given the analyses of Wimsatt[9] and/or Cartwright[10], *all* general statements that scientists use are best characterized as heuristics, or ceteris paribus laws. For instance, Cartwright[11] shows that "'fundamental" laws, like Newton's Law of Gravitation or Coulomb's Law, are false unless they are read as ceteris paribus laws. These "laws" work in many (even most) contexts, but in some contexts they are insufficient (and, strictly speaking, wrong) and other, perhaps less general, heuristics must be put to use.

---

[7]Bas C. Van Fraassen, *Scientific Representation: Paradoxes of Perspective* (Oxford: Oxford University Press, 2010).

[8]The fact that the content of the bottle plays a very important role in determining the adequate description of the content of the bottle means that it is very hard to imagine a context in which the content of the bottle is adequately described as marmalade or plastic.

[9]Wimsatt, *Re-Engineering Philosophy for Limited Beings : Piecewise Approximations to Reality.*

[10]Cartwright, *How the Laws of Physics Lie.*

[11]Nancy Cartwright, *The Dappled World: A Study of the Boundaries of Science* (Cambridge, UK : Cambridge University Press, 2003).

Following this discussion, the metaphysician might grant me that "water is H_2O" and other similar statements are not exact truths. In return, I have to recognize that it is a very good approximation to a general truth. So perhaps the mistake the metaphysician is making is not all that important?

I have no problem granting the strength of the approximation, but as I will argue in the next section, it makes quite a big difference for the possibilities of a certain type of metaphysics whether we treat statements like "water is H_2O" as truths or as very useful approximations. This is especially so if we simultaneously accept that we humans have limited cognitive capabilities.

4. Metaphysics as the search for fundamental truths

In the following sections, I will discuss the general view that a certain type of philosophy, usually referred to as metaphysics, should aim to provide us with a fundamental framework that can be used to understand all knowledge possessed by humans, and that this framework should be developed through the study of the most general features of reality.

There are many disagreements about whether it is possible, and, if so, how to define the methods and subject matter of metaphysics.[12] My aim here is not to give a general definition of metaphysics nor to dismiss the field as a whole. My aim is to challenge a certain vision for a type of philosophy that happens to be called metaphysics by those in favor of it. The question I raise in this essay is what consequences the view of scientific knowledge that I have sketched above, has for this specific kind of metaphysics.

An important first thing to note about this philosophical practice is that the aim of the practice is widely accepted to be unattainable. For instance, Gracia argues that "[m]etaphysics is an ongoing and never-ending enterprise, with no claim to finality".[13] Thus, when considering the value of such a practice, we must consider the costs and benefits in

[12]Jorge J. E. Gracia, *Metaphysics and its Task: The Search for the Categorical Foundation of Knowledge*. (Albany: SUNY Press, 1999).
[13]Ibid., p. 158.

relation to the pursuit of an unattainable goal. This means that what we need to consider is not what we would gain if we attained the goal, but rather the benefits of pursuing this goal.

Of course, we cannot consider this unless we further specify what this process consists in. I will therefore briefly consider two contemporary accounts of how such an all-encompassing system of knowledge propositions should be constructed, namely the accounts by Kit Fine[14] and Jorge Gracia[15] . Common to these accounts is that they start from a study of the most general and proceed down through the layers of generality using logic as an essential tool.

4.1. Contemporary Examples

Fine's discussion of the aim and method of metaphysics rests on the concept of *eidictic* statements. Eidictic statements concern the nature of reality. Fine's paradigmatic example is (of course) "water is H_2O"[16]. A statement is *eidictic as to content* if it *explicitly* refers to the nature of reality, e.g. "water is by its nature H_2O". Fine now gives the following description of the aim of metaphysics:

> Metaphysics should attempt to provide a foundation for all truths eidictic as to content; and what [...] provides the foundation are the metaphysical truths eidictic as to content along with [...] other possible truths that are not eidictic as to content. Thus given the non-edictic truths, the edictic truths of metaphysics will provide a foundation for all other edictic truths [17]

Gracia sees this project as too limited, and argues that metaphysics has been, and should continue to be an even broader study of the most general *categories,* and on the relation between less general categories and these most general categories. In this way, according to Gracia, metaphysics will attempt to build a foundation of all human knowledge[18]. Common to the two projects is that they focus on the

[14]Fine, *What is Metaphysics?*
[15]Gracia, *Metaphysics and its Task: The Search for the Categorical Foundation of Knowledge.*
[16]Fine, *What is Metaphysics?*, p. 10.
[17]Ibid., p. 11
[18]Gracia, *Metaphysics and its Task: The Search for the Categorical Foundation of Knowledge*, p. 220.

establishment of logical relations among *propositions* assumed to be *true*. This is obviously so in the project described by Fine, but also implicitly so in Gracia's description, to the extent that he describes metaphysics as investigating the logical relations between definitions of categories. If such relations are to establish a *foundation* of our knowledge, it seems crucial that the definitions of both the most general categories and at least some of the less general categories are true propositions.

4.2. Lessons from the Philosophy of Science

Two lessons from the philosophy of science indicate that these projects are far too ambitious. First of all, many philosophers now hold the so called *semantic view* of scientific theories.[19,20] This view denies that scientific theories are sets of propositions connected through logic. Rather, scientific theories should be viewed as sets of abstract models. This means that if all of our current scientific theories are to be fitted into a vast network of propositions, the metaphysician is faced with a tremendous translational task. Even if such translation is possible, the ambitious metaphysician is still faced with a second challenge. As argued in the first three sections of this essay, it is too simplified to view such general claims about nature as simple truths, they must be considered as context-dependent, *ceteris paribus* truths. Thus, a metaphysical system consisting of presumed true propositions will only give us a very inexact picture of our scientific knowledge and everyday knowledge.

A metaphysical system that relies so heavily on approximations will never be able to serve as a foundation of our current knowledge. Thus, if metaphysics is the search for a foundation of all knowledge, as both Fine and Gracia define it to be, we should consider finding a new methodology, since the one outlined so far cannot provide the foundation sought after. However, there may be another important

[19]Bas C. Van Fraassen, *The Scientific Image* (Oxford: Clarendon, 1980).
[20]Ronald N. Giere, *Explaining Science: A Cognitive Approach* (Chicago: Univ. of Chicago Press, 1988).

reason why we may still want to pursue the construction of this kind of metaphysical system.

Approximations are not problematic in themselves. Approximations can be extremely useful if they help us to *understand* the world we live in. Thus, if we have reason to believe that we might gain some additional understanding from the construction of a grand metaphysical system, we could consider whether the necessary approximations would be justified. Fine does not seem optimistic about this matter. The idea of a metaphysical explanation is, for him, "somewhat obscure"[21]. He suspects that simply a logical relation, as opposed to an explanatory relation, between the various eidictic truths will be sufficient for the aim of metaphysics. Similarly Gracia does not state that anything more than a logical connection between categories is required in order for us to have achieved a successful metaphysical system.[22]

In the philosophy of science literature, we find arguments that indicate that there is a limit to the extent to which complex logical systems can improve our understanding of the world. These arguments support Fine's doubt whether a metaphysical system will explain anything.

In relation to discussions on the philosophical interpretations of quantum physics, James Cushing [23] argued that humans are only capable of understanding certain types of explanations. Thus, in some cases there might be a qualitative difference between the best *description* (in the sense that it has the highest degree of empirical

[21]Fine, *What is Metaphysics?*, p. 11.

[22]These thoughts seem to be compatible with a unificationist view of explanation. See for instance: Michael Friedman, "Explanation and Scientific Understanding," *The Journal of Philosophy* 71, no. 1 (1974), 5-19.; or Philip Kitcher and Wesley C. Salmon, *Scientific Explanation* (Minneapolis: University of Minnesota Press, 1989). According to this view, we explain something by subsuming it under a general framework that encompasses many other things that were in need of explanation. So maybe the idea of constructing such a system in order to explain and understand the world we live in is not so "obscure" after all.The unificationist view of explanation has received severe criticisms over the past years (e.g. James Woodward, *Making Things Happen: A Theory of Causal Explanation* (New York: Oxford University Press, 2003)., Ch 8), but I do not think this criticism shows that it would be *impossible* to gain genuine understanding from the metaphysics outlines by Fine and Gracia. What the explanation literature certainly has taught us is that a unified view of nature is not *necessary* for our understanding of it.

[23]James T. Cushing, "Quantum Theory and Explanatory Discourse: Endgame for Understanding?" *Philosophy of Science* 58, no. 3 (1991), 337-358.

adequacy) of a class of phenomena – for instance radioactive decay or the formation of matter – and the best *explanation* of these phenomena (in the sense that it provides the most understanding of these phenomena).[24] Quantum physics probably provides the best examples of this. The Standard Model, for instance, allows physicists to predict the properties of all known elementary particles with incredible precision. If there has ever been a scientific theory that is empirically adequate, it is the Standard Model. Yet, for all its predictive precision many feel that the theory provides little understanding, partly because it has very counter-intuitive elements. Hence, physics lecturers often quote Richard Friedman[25] to the students who want to understand quantum mechanics: "Shut up and calculate!"

My point here is not that we should not try to understand quantum mechanics. My point is that if one is developing a metaphysical system *for the purpose of gaining understanding* – remember we gave up finding a logical foundation for our scientific knowledge when approximating *ceteris paribus* laws with true propositions - it may not be enough to require that we can make logical connections between the most general and the most particular, since this will not ensure that understanding is provided.

The reason for this is to be found in the human mind[26]. As a species, we are well adapted to survive in a world filled with objects behaving according to classical physics.[27] Would it be surprising if this ability came at a small cost, for instance, that we are unable to truly

[24]This argument put limitations on claims made by proponents of the epistemic theory referred to as 'inference to the best explanation' (see for instance Peter Lipton, *Inference to the Best Explanation*, 2nd ed. (London: Routledge/Taylor and Francis Group, 2004)). Though the heuristic that the loveliest explanations (i.e. those that provide the most understanding) is also the likeliest (as in likeliest to be true) might be very efficient in general, it is important to know that this does not apply to explanations of the very small or the very big.

[25] Or is it David Mermin? N. D. Mermin, "Could Feynman have said this?" *Physics Today* 57, no. 5 (2004).

[26]For further perspectives on the link between explanation and human cognition see Frank C. Keil and Robert A. Wilson, eds., *Explanation and Cognition* (Cambridge, Massachusetts: The MIT Press, 2000).

[27]Alison Gopnik, "Explanation as Orgasm and the Drive for Causal Knowledge: The Function, Evolution, and Phenomenology of the Theory Formation System," in *Explanation and Cognition*, eds. Frank C. Keil and Robert A. Wilson (Cambridge, Massachusetts: The MIT Press, 2000), p. 299-325.

grasp the behavior of elementary particles and black holes, unless we were able to talk about them in ways that are sufficiently similar to the ones we normally use?

4.3. Not suitable for fallible beings

What I have argued so far does not show that the projects outlined by Fine and Gracia could not possibly increase our understanding of the world we live in. However, my argument does show that if the projects are to increase our understanding, we may well have to require more than simply a logical relation among presumed truths to achieve it.

As we saw, the problem is that humans are limited beings, which means that the theories and explanations that can benefit us are not only constrained by the world around us, but also by our own cognitive capabilities.

Humans are limited beings in another respect as well: we are fallible in our reasoning. This basic condition of human existence provides us with an additional reason why we should consider choosing a different approach to metaphysics than the ones outlined by Fine and Gracia.

Consider a fallible metaphysician who wishes to argue for p on the basis of either one of the supposedly true premises q and r. Arguing from q to p involves n steps, while arguing from r to p involves n+m steps (m>0). Since the metaphysician is fallible, there is a non-zero chance of him making a mistake for each step he makes. Now, in the simplest model, where the probability of making a mistake during each step is constant, it is obviously preferable for the logician to argue from q and not r. In general,[28] the result is the same: "fallible thinkers should avoid long serial chains of reasoning".[29]

The metaphysician may grant me this and accept that metaphysics is very difficult. He could then object that the argument does not show that the top-down approach that I have discussed so far is more difficult than a more bottom-up approach that starts from less general knowledge and tries to abstract more general knowledge from it. Arguing for this,

[28]Wimsatt, *Re-Engineering Philosophy for Limited Beings : Piecewise Approximations to Reality*, Ch. 4.
[29]Ibid. p. 49

he could say that there is no a priori reason to think that it will take fewer steps to go from the most general to the more particular than the other way. This would be true, except that in the top-down approach one would have to argue for what the most general is, which, by the way, is what many of the metaphysicians of the past two and a half millennia have argued over.

Furthermore, fallible beings should reconsider constructing a system of knowledge from the top down because the strength of such a system will be highly dependent on the strength of a few general claims and the claimed connections between these claims and less general claims. This makes the system extremely fragile. Should one of these general claims turn out to be wrong, the consequences for the rest of the system would be devastating. A system of propositions that is built *from the bottom up* will be much less vulnerable when errors occur (see Figure 1).

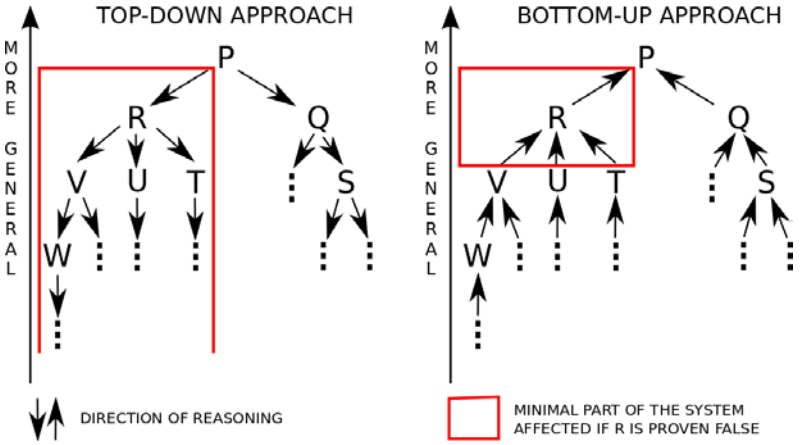

Figure 1: Two systems of propositions, constructed through different approaches. The advantage of the bottom-up approach is that it makes the system much more robust.

Thus, a fallible being who wants to construct a strong system of knowledge should (*ceteris paribus*) start from the less general and work his way up towards the most general instead of starting from the most general and work his way down. Such a bottom-up approach would not

only rely less on the truth of single propositions for its strength, it would also have other advantages, to which I will now turn.

5. Metaphysics for limited beings

No science is free of what we might call metaphysical assumptions— i.e. assumptions that are presupposed in the empirical exploration of nature. Scientists and other members of society recognize that there is a need to explicate these assumptions and to discuss their advantages and limitations. There is no question that philosophy can and should be an important voice in such discussions. However, the recognition that there is a need for philosophy does not mean that all kinds of philosophy are equally valuable. So far, I have characterized a particular kind of metaphysics – and thus a kind of philosophy – that seeks to create a unified system of presumed true propositions that encompasses all human knowledge by starting from the most general and working its way down to the most particular. We have seen that this approach to philosophy has severe limitations. Choosing a more practice-oriented and in a sense more bottom-up approach to philosophy would increase the robustness of the system that is constructed, but it would also allow the philosopher to construct a view of current science that distorts the products of science to a lesser extent and thus provides a better understanding of these products.

Take causation as an example. This is certainly a general category or concept that is used in many sciences, and sometimes in counterproductive ways. A philosophical analysis of causation will definitely be valuable in its own right and as an input to various discussions in science and society. My point here is that such an analysis is more fruitful if it starts from the actual practices of speaking about and working with causation. Understanding (or improving) these complex practices will require the use of abstractions and approximations. But our understanding need not, and should not, start from one general definition of what causation is and how it ought to be used. We will be much better suited for choosing the fruitful and appropriate approximations if we start by

considering the actual diverse practices and their aims. Through this approach we might be able, with time and effort, to understand how causation is used in, for instance, biology or Patagonian literature. We may also come to understand why it used in these ways in the given practices and attempt to evaluate this usage given an analysis of the aims of these practices. We might even try to understand the relation between causation in biology and Patagonian literature. I take it to be an empirical question whether or not such a comparison would show significant similarities. It might turn out that they have so little in common that it makes little sense to say that causation means the same thing in the two practices. However, this does not mean that the analysis of causation in biology or Patagonian literature is worthless, or that we cannot achieve a unified account of causation in, say, biology and chemistry. What it does mean is that a definition of causation that captures every use of the concept is unachievable.

More generally: What a philosopher aiming to understand the relations between different domains of knowledge can do is to try and connect the clusters of heuristics used in various practices into new and perhaps bigger clusters. In some cases, the connection formed in these new clusters may be strictly logical, but disclosing other, or perhaps additional, relations, such as various explanatory relations, may provide important new understanding as well. In any case, it is important to remember the limitations of the heuristics when seeking connections between them in order to avoid overly general conclusions. This means that in order to succeed, the philosopher must have knowledge about the scope and limitations of the heuristics she is trying to unify, or at least be involved in a continuous discussion with people who have.

Such a practice-oriented approach can be pursued to gain a deeper understanding of the practice with minimal distortion in the description of the tools and products of the practice. This approach will preserve a nuanced picture of our knowledge and keep us aware of the limitations to our general conclusions. Furthermore, it will help us to stay within our cognitive limits. Each step will be potentially interesting and useful for scientists who are usually more interested in the specific historical, social and

logical connections between their own field and other specific fields rather than a very general picture of how all knowledge can be connected; this is especially true if the more general picture comes at the cost of significant distortion of the practice of each field. Thus, this approach will be better suited to the task of answering the questions from scientists that Sullivan & Pannier[30] state as one of the reasons why metaphysics is still important.

In short, a practice oriented, more bottom-up approach to metaphysics has the same potential, as a top-down approach, but is far less vulnerable and has more immediate relevance for discussions outside philosophy and is thus more suited for limited beings like human philosophers.

6. Conclusions

Before his analysis of the detailed characteristics of metaphysics, Fine remarks that:

> Metaphysics [...] might be a somewhat anemic discipline – there might be very little for it to do. But it is also thought that metaphysics might play an important foundational role. It is not merely one form of enquiry among others, but one that is capable of providing some kind of basis or underpinning for other forms of enquiry.[31]

In his conclusion, Fine further reflects on the possibility of metaphysics, and concludes that it is not unreasonable to hope that we may one day find the truths that underpin all others, and thus that the task of metaphysics is worth pursuing.

My aim in this essay has been to argue that even though Fine may be right in his conclusion that metaphysics, as he describes it, is not impossible, it will not serve the foundational role that he and others want it to serve. This leads us to ask what other aims the methodology of metaphysics might serve and whether it is the most suited methodology for pursuing these aims. One important alternative aim is to increase

[30]Thomas D. Sullivan and Russell Pannier, "The Bounds of Metaphysics," in: *What are we to Understand Gracia to Mean?: Realist Challenges to Metaphysical Neutralism*, ed. Robert A. Delfino (Kenilworth : Rodopi, 2006), p. 1-12.
[31]Fine, *What is Metaphysics?*, p. 11

our understanding of the structure of our knowledge and of nature in general. In this essay, I have presented a number of reasons why I think a more bottom-up and practice-oriented approach to the study of scientific and everyday knowledge will be better suited for the philosopher pursuing this aim than the top-down approach that is often associated with the practice of metaphysics.[32]

Literature

Cartwright, Nancy. *The Dappled World: A Study of the Boundaries of Science*. Cambridge, UK : Cambridge University Press, 2003.

———. *How the Laws of Physics Lie*. Repr. ed. Oxford: Clarendon, 1984.

Chang, Hasok. *Is Water H_2O? : Evidence, Realism and Pluralism*. Boston Studies in the Philosophy of Science Ser. New York : Springer, 2012.

Cushing, James T. "Quantum Theory and Explanatory Discourse: Endgame for Understanding?" *Philosophy of Science* 58, no. 3 (1991): 337-358.

Fine, Kit. "What is Metaphysics?" In *Contemporary Aristotelian Metaphysics*, edited by Tahko, Tuomas E., 8-25. Cambridge: Cambridge University Press, 2012.

Friedman, Michael. "Explanation and Scientific Understanding." *The Journal of Philosophy* 71, no. 1 (1974): 5-19.

Giere, Ronald N. *Explaining Science: A Cognitive Approach*. Chicago: Univ. of Chicago Press, 1988.

Gopnik, Alison. "Explanation as Orgasm and the Drive for Causal Knowledge: The Function, Evolution, and Phenomenology of the Theory Formation System." In *Explanation and Cognition*, edited by Keil, Frank C. and Robert A. Wilson, 299-325. Cambridge, Massachusetts: The MIT Press, 2000.

Gracia, Jorge J. E. *Metaphysics and its Task: The Search for the Categorical Foundation of Knowledge*. Albany: SUNY Press, 1999.

Keil, Frank C. and Robert A. Wilson, eds. *Explanation and Cognition*. Cambridge, Massachusetts: The MIT Press, 2000.

Kitcher, Philip and Wesley C. Salmon. *Scientific Explanation*. Minnesota Studies in the Philosophy of Science ; 13. Minneapolis: University of Minnesota Press, 1989.

Kripke, Saul A. *Naming and Necessity*. Reprint ed. Oxford: Blackwell, 1988.

Lipton, Peter. *Inference to the Best Explanation*. International Library of Philosophy. 2nd ed. London: Routledge/Taylor and Francis Group, 2004.

Mermin, N. D. "Could Feynman have Said this?" *Physics Today* 57, no. 5 (2004).

[32] I wish to thank everyone involved in making the Summer School an inspiring and memorable experience, especially Sebastian Krebs, Simon Baumgartner and Thimo Heisenberg. You did a wonderful job! This essay benefitted greatly from critical comments from Shyane Siriwardena and from the graphical skills of Natalia Golubeva.

Putnam, Hilary. "Meaning and Reference." *The Journal of Philosophy* 70, no. 19, Seventieth Annual Meeting of the American Philosophical Association Eastern Division (Nov. 8, 1973): pp. 699-711.

Sullivan, Thomas D. and Russell Pannier. "The Bounds of Metaphysics." In *What are we to Understand Gracia to Mean? : Realist Challenges to Metaphysical Neutralism*, edited by Delfino, Robert A., 1-12. Kenilworth : Rodopi, 2006.

Van Fraassen, Bas C. *The Scientific Image*. Oxford: Clarendon, 1980.

———. *Scientific Representation: Paradoxes of Perspective*. Oxford: Oxford University Press, 2010.

Wimsatt, William C. *Re-Engineering Philosophy for Limited Beings : Piecewise Approximations to Reality*. Cambridge, MA: Harvard University Press, 2007.

Woodward, James. *Making Things Happen: A Theory of Causal Explanation*. Oxford Studies in Philosophy of Science. New York: Oxford University Press, 2003.

Understanding the Dispute between Science and Metaphysics

Shao-Pu Kang
(National Taiwan University)

Modern history is marked by the rapid development of science and technology. In such a scientific era, metaphysics is often seen as an anti-modern discipline of the past. The logical positivists even once claimed that metaphysics should be eliminated from philosophy in order to save the genuineness of philosophy. However, does science really undermine the credentials of metaphysics and render it redundant? This paper dismisses the dispute between science and metaphysics as a pseudo debate. It characterizes the dispute as one between one totality of statements, labeled science, and another totality of statements, label metaphysics. By showing that metaphysics has a non-empirical feature, which science lacks, it argues that metaphysical statements and scientific statements do not belong to the same kind, and *ipso facto* there is no common base for comparison. Hence, the legitimacy of metaphysics is unaffected, however successful science might be. In the penultimate section, it sketches two positive contributions metaphysics might have for science: first, metaphysics can explicate the concepts used in science; second, metaphysics can delimit the boundary of science. This paper does not aim to propose that metaphysics should govern scientific research, but rather to show that both science and metaphysics are legitimate fields of rational enquiry worthy of equal respect.

1. Introduction

Kant cast doubt on the possibility of pursuing a metaphysics whose aim is, as traditionally conceived, to tell us the fundamental structure of reality as a whole.[1] For him, the limit of human understanding prevents us from gaining knowledge of things-in-themselves. Kant argued that metaphysics should confine itself to the fundamental structure of our *thought* about reality, not to the structure of reality itself. The anti-metaphysical trend of thought culminated in the works of the logical positivists, the most vigorous opponents of metaphysics in the twentieth

[1] Immanuel Kant, *Critique of Pure Reason*, trans. Kemp Smith (1787; repr. London: Macmillan, 1929), 56-57.

century. Ayer set up the criterion of verifiability, in terms of which a sentence is said to be factually significant if and only if it is verifiable in principle. Since no observations are relevant to determine the truth or falsehood of metaphysical statements, these, according to this criterion, are nonsensical and should thus be eliminated from philosophy in order to save its genuineness.[2]

Although the attack of logical positivism proved unsuccessful, metaphysics was far from being a legitimate field of rational enquiry respected by all. The twentieth century was an era characterized by the rapid development of science and technology, which dramatically changed the ways of life and the general form of culture among human beings. Comparatively, metaphysics hardly produced anything capable of increasing convenience and efficiency. This difference between science and metaphysics made the latter bear an image of uselessness and obsolescence. However, is metaphysics really a field of enquiry belonging only to the past? Should philosophy distance itself from metaphysics in order to be a genuine branch of knowledge? Or should metaphysics be eliminated due to the success of science?

While the status of metaphysics might be challenged from different points of view, this paper focuses only on those challenges which arise owing to science's great success, which metaphysics allegedly lacks. They might argue that in such an era marked by the advance of science, metaphysics becomes a field of enquiry which is not worth undertaking anymore. The more radical ones in the scientific camp might even contend that science can answer any question, and whatever science cannot answer is not a genuine question at all. Indeed, human intellectual history reveals a bitter past for philosophy, in the course of which various disciplines separated themselves from philosophy and became independent. It embarrassed philosophers even more when those fields of enquiry, which were no longer under the reign of philosophy, experienced a great breakthrough in the past four hundred years. However, does science really undermine the credentials of metaphysics and render it redundant? Can science replace metaphysics and take over the job to deal with such a question as what the nature of

[2] Alfred J. Ayer, *Language, Truth, and Logic* (New York: Dover Publications, 1952), 13-29.

reality is? Before answering any similar question, we must first understand the nature of the dispute between science and metaphysics.

2. The Nature of the Dispute

It is vague to say that science undermines the credentials of metaphysics, or that there is a dispute between them. To understand the dispute in question, we need to have a clearer idea about exactly (1) *what* disputes with *what*, (2) *why* there is such a dispute, and (3) *in what sense* one undermines the other. What I have in mind concerning the first question is that the dispute is between two sets of statements: scientific statements and metaphysical statements. Specifically, it is a dispute between one totality of statements, labeled science, and the other totality of statements, labeled metaphysics, not one between terms or individual statements taken in isolation.[3] Statements labeled science include mathematical formulae and models, the laws of physics, and assertions in the special sciences;[4] statements labeled metaphysics include statements dealing with universal, causation, and other topics often regarded as metaphysical ones. One should note that I am neither defining what a scientific (metaphysical) statement is, nor am I saying that all statements labeled science (metaphysics) share certain features necessary for being a scientific (metaphysical) statement. In my characterization, statements are scientific ones simply because they occur in disciplines generally recognized as science such as physics, biology, or psychology, or because they are asserted by those who are deemed scientists. By the same token, statements are metaphysical ones because they discuss issues often included in metaphysical textbooks, such as necessity, existence, or identity, or because they are asserted by those who claim to be metaphysicians. It is in this sense that I call it a dispute between one totality of statements labeled science and the other totality of statements labeled metaphysics.

[3] Hereafter I use "scientific (metaphysical) statements" and "the totality of statements labeled science (metaphysics)" interchangeably.

[4] One might argue that mathematical formulae and models account for the bulk of science, although they are not statements at all. Here I use the word "statement" in a rather broad sense only for the sake of argument.

Why is there such a dispute between science and metaphysics? Presumably, the dispute results from the alleged fact that both science and metaphysics assert something about the nature of reality in one way or another. For instance, physics tells us what the basic particles constituting reality are, biology tells us about the compositions and functions of various organisms, and psychology explains the relationship between human minds and the corresponding neural processes. On the other hand, metaphysics makes assertions such as: there are universals, physical things and physical events exhaust the whole reality, or free will is compatible with determinism. What follows naturally from this alleged fact is that people would compare science and metaphysics according to different criteria.

There are various ways of comparing two systems of statements, e.g. we can compare them in terms of the degree of inner consistency, of the scope of explanation, or of simplicity. Arguably, the difference between science and metaphysics is marked predominantly by the fact that science has practical value while metaphysics lacks this. Science has practical value, that is, the collective work of the statements labeled science can produce concrete things existing in time and space, and the systematic employment of those things can make human life more efficient and convenient; metaphysics, on the contrary, enjoys no such practical value.[5] It is in this sense that science undermines the credentials of metaphysics. Now we understand the dispute between science and metaphysics as one between one totality of statements labeled science and the other totality of statements labeled metaphysics; science challenges the legitimacy of metaphysics, and whereas both manage to probe into the nature of reality in one way or another, only science possesses practical value.

[5] I am not defining what it takes to have practical value; such words as "systematic," "efficient," or "convenient" are too vague to be constituent terms of a definition at all. That I characterize practical value in this way is only for the sake of discussion. In what follows I use "practical," "useful" and "pragmatic" interchangeably.

3. Dismissing the Dispute

Science can undermine the credentials of metaphysics only if the dispute in question is a genuine debate. And a dispute between two systems of statements would not be a genuine one unless they both consisted of the same kind of statements. For instance, we do not usually compare moral judgments with scientific statements, since each asserts things which are quite unlike that which the other asserts. While science is concerned with what *is*, ethics copes with what *ought* to be, and that there is a gap between "is" and "ought" is widely acknowledged, if not accepted. The point is that if two sets of statements are not of the same kind, we cannot compare them, since there is no common basis for comparison at all. If that is the case, then for there to be a genuine dispute between science and metaphysics, one must presuppose that scientific statements and metaphysical statements belong to the same kind. I doubt the plausibility of this presupposition. Before entering my criticisms of this presupposition, let me briefly explain the strategy underlying my argument. Neither do I believe that there is a general and unambiguous criterion which enables us to draw a sharp dividing line between science and metaphysics, nor do I intend to define what it takes for a statement to be a scientific or metaphysical statement, and which would then allow a classification. What I do believe is, if scientific statements and metaphysical statements belong to the same kind, they must share the same features to at least a sufficiently high degree. If, however, we find some apparent features that scientific statements have, which metaphysical statements lack, or vice versa, then we have good reason to believe that scientific statements and metaphysical statements are not of the same kind. We are justified in doing so even without defining beforehand exactly what a scientific or metaphysical statement is, or what we mean with the word "kind". Hence, in what follows I am going to show some key features that science and metaphysics do not share.

A particular science is supposed to deal with the nature of being of a particular kind, e.g. biology concerns the nature of biological organisms, chemistry the chemical elements, and physics the basic particles. That is, these particular sciences all purport to explore one certain part of reality. Metaphysics, however, concerns the nature of being *as a whole*.

Metaphysics is recognized as the science of "being *qua* being", in the sense that it is the systematic study of the most fundamental structure of reality as a whole.[6] The fact that two sets of statements assert something different might not be sufficient to separate scientific statements and metaphysical statements into two kinds. It is the "non-empirical" character of metaphysics – metaphysics does not typically appeal to experimental or observational evidence in support of its claims – that makes it distinct from science. For instance, in claiming that there are universals, it would be fruitless to try to confirm it by collecting empirical data, nor can it be refuted simply because all things observed are particular objects. Such a claim must be justified or rejected only by sound metaphysical arguments. This echoes what Fine believes to be one of the five features, which, when combined, serve to distinguish metaphysics from other forms of enquiry, that is, the "aprioricity" of its method.[7] As Fine states: "The claims of science rest on observation; the claims of metaphysics do not, except perhaps incidentally. Its findings issue from the study rather than from the laboratory."[8]

The fact that metaphysical statements are non-empirical, while scientific statements are empirical, makes their truth-makers and the ways of their truth-making different. The truth-makers of scientific statements are facts. It is the nature of mind-independent reality that renders scientific statements true or false. And scientific statements are true by correctly representing the actual states of affairs. For instance, the statement "metal expands when heated" is true in virtue of its correctly representing the fact that metal expands when heated. The truth-makers of metaphysical statements, on the contrary, are arguments. Metaphysical truths are not, except incidentally, determined by empirical enquiry of whatever kind, but by arguments metaphysicians give. And metaphysical statements are true if they

[6] E. Jonathan. Lowe, *The Possibility of Metaphysics: Substance, Identity, and Time* (New York: Oxford University Press, 1988), 1-27. See also Lowe, *A Survey of Metaphysics* (New York: Oxford University Press, 2002), 1-20.

[7] Kit Fine, "What Is Metaphysics," in *Contemporary Aristotelian Metaphysics*, ed. Tuomas E. Tahko (Cambridge: Cambridge University Press, 2011), 8-25. The five features are the aprioricity of the methods of metaphysics, the generality of its subject-matter, the transparency or "non-opacity" of its concepts, its eidicity or concern with the nature of things, and its role as a foundation for what there is.

[8] Ibid., 9.

logically follow from some finite and logically consistent class of metaphysical arguments. For instance, the proposition "the only things that exist are physical objects and physical events" is true since physicalism, let us say, logically follows from some logically consistent class of metaphysical arguments, while mind-body dualism or idealism does not.

Now we get a clearer picture on how metaphysics differs from science: metaphysial statements have a non-empirical feature which scientific statements lack, and *ipso facto* both their truth-makers and the ways of their truth-making are different. As I have said, two sets of statements are comparable only if they belong to the same kind. Though both concern the nature of reality, scientific statements and metaphysical statements are not comparable, since the targeted fields of science are certain parts of reality, while metaphysics aims to illustrate the nature of reality taken as a whole. Science makes claims about the empirical world, while metaphysics does not aim to represent the empirical world which is, in principle, if not in practice, observable through sense-perception. Scientific statements and metaphysical statements have their own truth-makers and their specific ways of truth-making, and the truth-makers of scientific statements cannot render metaphysical statements true or false, and vice versa. The conclusion we arrive at here is, scientific statements and metaphysical statements are not of the same kind, and thus the dispute between them is not a genuine one. By this very fact, the credentials of metaphysics is unaffected by the great success that science has achieved.

4. Possible Objections

Objection 1: The usefulness of science is one of the significant marks which represent a higher likelihood of its being true; by contrast, the truth of metaphysical statements is far less probable due to their lack of practical value. Owing to this huge difference in the likelihood of being true, science still undermines the status of metaphysics, irrespective of whether they belong to the same kind. This approach, however, has two problems. First, usefulness does not entail truth. This has been shown clearly in the history of science. For instance, Newtonian mechanics is

still widely applied in mechanical engineering, though it is false, strictly speaking, whereas quantum mechanics is difficult to apply in ordinary cases, though it is the best theory for the time being. The problem is that being useful is not constitutive for a sentence to be true; there is no necessary logical connection linking usefulness and truth together. People might tend to accept sets of statements that are more useful and have less confidence in sets that are not, but that is only a practical consideration which by no means entails that the sets accepted are true. Second, science manifests its practical value only by the collective work of the statements labeled science. In other words, a statement taken in isolation would generate nothing; only when combined together to form a corporate body can the totality of scientific statements produce anything useful. Specifically, in a scientific theory there are often some highly formalized parts, e.g. mathematical formulae and models; those formalized parts would not have any empirical bearing unless they receive some empirical interpretations, that is, some bridge principles tell us which variables in the formulae stand for which objects observable in sense-perception. This empirical interpretation would render the formalized parts relevant to the phenomena of our experience, and it is in this sense only that the collective work of the formalized parts and the empirical parts can generate concrete things that advance the well-being of human beings. The point I want to emphasize is, that while individual statements might be bearers of truth, only through the cooperation of various statements working as a corporate body can science have practical value. This shows that usefulness is irrelevant to truth: the usefulness of a totality of statements depends on what the constituent statements are and the way they are combined together; truth, however, is determined by whether what is asserted in a given statement obtains.

Objection 2: Practical considerations would dominate the process of theory choice when it is underdetermined by all given evidence. At any given time there might be two or more empirically adequate theories that are compatible with the same observable phenomena, while conflicting with each other in their fundamental structures. Under those circumstances, evidence favors none of the empirically equivalent theories, that is, theory choice is underdetermined by purely epistemic considerations; we cannot but, among many theories, choose one

according to some pragmatic considerations.[9] Based on this understanding, one might thereby argue that practical value alone would eventually prompt people to accept science rather than metaphysics. In this sense science still undermines the legitimacy of metaphysics.

However, even when what is at issue is narrowed down to practical value, science still fails to challenge the status of metaphysics. For there to be a situation underdetermined by evidential considerations, it is necessary that all rival theories are *scientific* theories. Scientific theories are incommensurable only with scientific theories, if any. A scientific theory does not compete with a metaphysical theory, since scientific statements and metaphysical statements are not of the same kind. This means we would not encounter such a situation in which theory choice between science and metaphysics is underdetermined by evidence and is governed rather by pragmatic considerations. Thus, it is futile to argue that science undermines the credentials of metaphysics by appealing to the underdetermination thesis.

5. What Can Metaphysics Contribute to Science ?

What has been argued in the previous part is merely a negative account; it only says that the credentials of metaphysics cannot be undermined by science, however successful it is. This is far from satisfactory for those determined disbelievers of metaphysics; they might think that even if science cannot reject metaphysics as such, metaphysics is completely irrelevant to scientific research and is thus totally insignificant. Before ending this paper, I would like to briefly sketch two positive contributions that metaphysics might have for science: (1) metaphysics can explicate the concepts used in science; and (2) metaphysics can delimit the boundary of science.

Let us take the issue of consciousness as an example to explain this. Cognitive neuroscientists study the micro-structure of the brain and its neural mechanisms correlating with various cognitive activities, and aim thus at a final account for consciousness in purely neural-biological

[9] Bas van Fraassen, *The Scientific Image* (Oxford: Clarendon Press, 1980), 87-96. This possible objection is formulated based on van Fraassen's constructive empiricism.

terms.[10] Now consider one thesis that is widely acknowledged: the phenomenal character of consciousness[11] supervenes on the neuro-biological functioning of organisms. What neuroscientists have in mind is that there is a correlation between, let us say, pain feeling and C-fiber stimulation. Metaphysicians might explicate the notion of supervenience in this way:

> *Pain-C-fiber Supervenience*. Pain feeling supervenes on C-fiber stimulation in that if *x* feels any pain, there is a stimulation of C-fiber such that *x*'s C-fiber is stimulated, and necessarily any object whose C-fiber is stimulated feels pain.[12]

Supervenience as such is a covariance relation between the supervenient and the supervened. It does not follow, however, that there is a causal relation. The existence of the supervened makes a causal difference, but that does not suffice to confer causal efficacy on the supervenient. Thus, the pain-C-fiber supervenience thesis only says that one necessarily feels pain whenever his C-fiber is stimulated, without claiming that it is C-fiber stimulating that causes pain. By explicating the concepts used in science in a similar manner, we can understand what scientists have and have not achieved in a much clearer way.

The first contribution metaphysics might have for science is actually related to the second one. If such a pain-C-fiber supervenience thesis does exist, metaphysicians might still ask the following questions: it happens to be the case that C-fiber stimulating correlates with pain, but why is this so? Why does C-fiber stimulating not correlate with itch but pain? This can be put more generally: why do conscious states correlate with those neural states with which they correlate? All questions of this and of similar kinds are actually asking why there are such supervenience relations between conscious states and neural states. The answers depend on how we interpret the word "necessary" in the supervenience thesis. If pain necessarily supervenes on C-fiber

---

[10] Michael S. Gazzaniga, Richard B. Ivry, and George R. Mangun, *Cognitive Neuroscience: The Biology of the Mind*, 3rd ed. (New York: Norton, c2009). My knowledge of cognitive neuroscience is based mainly on this book.

[11] Thomas Nagel, "What Is It Like to Be a Bat?" *Philosophical Review* 83 (1974): 435-50. The phenomenal character of consciousness refers to the *what-it-is-like* aspect of conscious experience coined by Nagel in his 1974 paper.

[12] Jaegwon Kim, *Philosophy of Mind*, 3rd ed. (Boulder, CO: Westview Press, c2011), 8-10. I revise Kim's formulation of supervenience developed in his 2011 book.

stimulating in the *nomological* sense, neuroscientists might well claim that it is the way the world is that renders the supervenience thesis true. However, it is quite reasonable to imagine that the way the world actually *is,* is not the only way the world *could* be. For example, I am actually writing this paper now, but I could have done otherwise; I might have felt tired of writing paper and started to play video games. Similarly, it might actually be the case that C-fiber stimulating correlates with pain, but it does not necessarily correlate with pain in the *metaphysical* sense. The fact that the human brain is constituted in a certain way is a contingent fact, so is the supervenience relation between C-fiber stimulating and pain, if any. Here we can understand what is meant by saying that metaphysics can delimit the boundary of science: scientists can only investigate the way the world actually is without explaining why it must be so or what may be, but happens not to be, the case; on the other hand, metaphysicians chart various metaphysical possibilities of reality, and tell us that the world we happen to live in is only one among many possible ones.

6. Conclusion

This paper characterized the dispute between science and metaphysics as a dispute between one totality of statements, labeled science, and the other totality of statements, labeled metaphysics. I argued that it is because scientific statements have practical value while metaphysical statements lack this, that science allegedly undermines the legitimacy of metaphysics. However, for there to be a genuine dispute between two systems of statements, they must belong to the same kind. In Section 3, I argued that the non-empirical feature of metaphysics makes it distinct from science. By showing that there is such a distinctive feature which one set of statements has, which the other set does not have, we are justified in believing that metaphysical statements and scientific statements do not belong to the same kind. This means, however, that there is no common basis for comparing science and metaphysics, and accordingly the dispute between them is not a genuine debate. Hence, the credentials of metaphysics are unaffected, however successful science has been. In Section 5, I sketched two positive contributions

metaphysics might have for science. On the one hand, metaphysics can explicate the concepts used in science; on the other hand, metaphysics can delimit the boundary of science. Above all, science and metaphysics are not mutually exclusive; both are legitimate fields of rational enquiry that deserve equal respect.

Literature

Ayer, Alfred J. *Language, Truth, and Logic*. New York: Dover Publications, 1952.

Fine, Kit. "What Is Metaphysics?" In *Contemporary Aristotelian Metaphysics*, 8-25. Edited by Tuomas E. Tahko. Cambridge: Cambridge University Press, 2011.

Gazzaniga, Michael, Richard B. Ivry, and George R. Mangun. *Cognitive Neuroscience: The Biology of the Mind*. 3rd edition. New York: Norton, c2009.

Kant, Immanuel. *Critique of Pure Reason*. Translated by Kemp Smith. 1787. Reprint, London: Macmillan, 1929.

Kim, Jaegwon. *Philosophy of Mind*. 3rd edition. Boulder, CO: Westview Press, c2011.

Lowe, E. Jonathan. *The Possibility of Metaphysics: Substance, Identity, and Time*. New York: Oxford University Press, 1998.

——— . *A Survey of Metaphysics*. New York: Oxford University Press, 2002.

Nagel, Thomas. "What Is It Like to Be a Bat?" *Philosophical Review* 83 (1974): 435-50.

Van Fraassen, Bas. C. *The Scientific Image*. Oxford: Clarendon Press, 1980.

Thirsting for a world: the permeating presence of metaphysics in Epistemology

Jesus Moreno
(Simon Fraser University)

Recent literature in epistemology seems to argue that epistemology is better carried out under the guise of an eliminative metaphysics, if not with the goal of eliminating the role of metaphysics in the field altogether. Normative disagreement, for example, has been presented as what can be better understood under the heading of conflict between choices. In this paper I will argue that, due to some basic but tremendously important phenomenological features of engaging in discourse with other people, any attempt to account for normative disagreement by representing it as conflict is condemned to an explanatory failure. Moreover, this phenomenology is indicative of a much deeper concern: one connected to an attitude that is metaphysical in nature. This attitude, namely our conception of a strong and substantial relation between ourselves as the kind of epistemic agents we are and the world as a knower-friendly place, plays a constitutive role in practices such as epistemology—something that is better evidenced once we analyze the nature of global skeptical arguments.

Introduction

To some, presupposing a substantial metaphysics in our epistemology is somewhat passé. Epistemology, so they seem to claim, is to be seen as a formal enterprise—one filled with barren models and entities the presence of which is adjudicated on the strictest notions of explanatory necessity. Debates such as that on the role and nature of normativity, and in particular of epistemic normativity, are to be carried out under the guise of an eliminative metaphysics. Normative disagreement, for example, has been presented as a mere instance of conflict between choices. In this paper I will argue that, due to some basic but tremendously important phenomenological features of engaging in discourse with other people, any attempt to account for normative disagreement by representing it as conflict is condemned to an explanatory failure. Whereas conflict needs not be factual, the

phenomenology of disagreement dictates for the intentions of the speakers to refer to the world—and thus, it is concerned with factual issues. Moreover, this phenomenology is indicative of a much deeper concern: one connected to an attitude that is metaphysical in nature. This attitude, namely our conception of a strong and substantial relation between ourselves as the kind of epistemic agents we are and the world as a knower-friendly place, plays a constitutive role in practices such as epistemology—something that is better evidenced once we analyze the nature of global skeptical arguments.

I.

Let us first introduce a basic framework within which we may discuss normative disagreement. By "Normativity" I shall refer to normativity in some realist sense—that is, a strong, robust, factual notion of normativity where normative demands stand independently of our attitude towards them. By "normativity*" I shall be referring to the weaker, deflationary-style notion of normativity discussed by non-realists accounts—a normativity the weight of which is derived from one form or another of relativization to goals. This entails that we cannot talk about norms irrespectively of the kind of theory discussed. Let us differentiate then between what we shall call "Norms" and what we shall call "norms*". Thus, a Norm would be a fact (e.g. you ought to do x), and it will play a crucial role in simpliciter ought-statement found in Normative accounts and to which an agent appeals when holding beliefs of the kind "given [facts, including the relevant normative fact], I ought to do x". In contrast, norms* are closer to policies (e.g. if you want to achieve y, then do x; or always believe v if you believe w), and beliefs containing norms* can be presented as holding something along the lines of the formulaic "given [facts], and policy ψ, I ought* do x".[1]

[1] Here I am appealing to something rather close to Hartry Field's notion of norms as policies in accordance to his presentation of it in his "Epistemology Without Metaphysics" (Field 2009). Field's assessor-relativist account embodies, and develops, several key notions from the non-realist side—notions that can be found in the works of authors such as MacFarlane and Gibbard, see: (MacFarlane, Making sense of relative truth 2005), (MacFarlane 2009), (Gibbard, Wise Choices, Apt Feelings 1990), (Gibbard 2003), (Gibbard 2008). For an analysis of reductivist and eliminativist accounts of normativity (such as

Given the difference between Norms and norms*, it is possible to see that not all normative claims are created equal. Normative claims that appeal to Norms appear to derive whatever normative weight they have from the factual nature of the Norm themselves—that is, it is that fact in the world that a Norm is that which what imposes a demand on you. Claims that employ norms*, however, seem to derive their normative weight from a conditional based as much on our desires as on the way the world is. Claims that appeal to Norms seem to believe in a strong, substantial normativity. This Normativity—whether existent or not—is qualitatively different from that which grounds claims that employ norms*. For the latter, "normativity" stands closer to the kind of guidance power offered by goal-oriented rules of thumb, by minimax strategies aiming to provide an algorithm with which to calculate the payoff from different paths of action. Unlike Normativity, this normativity* poses but a contingent demand upon us: "if the world is a certain way, and you *desire* a certain outcome, then this is the way to go". This would not be a problem for non-realists were their account to show that an appeal to normativity* was indeed simpler, stronger, and better fit for an understanding of normative disagreement than its aforementioned alternative. But despite their apparent parsimony, such accounts can hardly be seen as simpler and stronger without first effecting a massive misinterpretation of basic dialectical facts.

II.

Since norms* are policies, and policies are contingent in their relevance to at least one desired goal, a conflict between norms* is a conflict between desired outcomes. If I want to have Chinese food and my friend wants to have French food, then we are in a scenario where our desires are incompatible in such a way that only one of them, but not both, can be satisfied (assuming, of course, that we also want to have dinner together). Our desires, that is our preference as for where we are to dine, are in *conflict*. Were we to express our desires by saying "we should have

(Gibbard 2003)) that shares several of the worries raised in this essay see chapter 4 in Barry Stroud's *Engagement and Metaphysical Dissatisfaction* (Stroud, Engagement and Metaphysical Dissatisfaction 2011).

__ for dinner" we would then be employing the "should" of preference (i.e. "given [facts, including my desire to x] and [policy to satisfy one's desire], we should* do y).

The sort of phenomenon that we wish to represent under the notion of conflict can be better individuated via the following analogy: Think of a typewriter. The mechanism executing a key's command in a typewriter can be such that, when several keys are pressed simultaneously, the type bars end up entangled. That is to say, not all commands can be executed at once. The machine can successfully execute but one command at any given time. If more than one command is given at the same time (such as when we press several keys simultaneously), not all commands succeed. Were we to think of the pressing of a key as a preference, we would conceive of the situation as one of conflict among preferences: ceteris paribus, only one option can be carried out at any given time successfully—that is, only one preference can be satisfied at any given time.

This kind of incompatibility can also rise between agents' choices. If Smith and Jones come to a fork while hiking down a path, and each one favours the road rejected by the other, either one agent's choice will be ignored or they will have to travel separately. In some cases, the latter disjunct may not be a feasible option (imagine them driving down a highway instead). Their choices are incompatible. Whereas some would say that our travelers *disagree* on what path to take, a seldom few would wish to extend this claim to the typewriter scenario. After all, in what sense, if any, do the typebars *disagree*? It could be claimed that, unlike the typebars, our travellers can argue in favour of their choice—for example, by appealing to *instrumental normativity*. In a basic form, this scenario would go along the following lines:

S: We should go left
J: We should go right

Let us desensitize S and J the following way:

S*: Given the facts x, y, and z, and policy α, we should go left.
J*: Given the facts x, y, and z, and policy β, we should go right.

Assuming Smith accepts J* as true, and Jones accept S* as true (in other words, that each accepts the other's desensitized claim to be true), it appears to follow that we are in the presence of a conflict between Smith's choices and Jones' choices. There is no disagreement between S* and J*, and yet there was disagreement between S and J. How is this situation to be solved? Smith and Jones cannot *debate*, insofar as they each accept the truth of the other's desensitized sentence and, therefore, no argument can be made against the truth of either's assertion. However, they can *negotiate*. In negotiation, agents aim at maximizing the degree to which their individual choices will be satisfied. In a debate, agents are concerned with the truth-value of a claim—a common understanding of which is a must if the agents do not wish to talk across each other. When two people debate whether a sentence is true, no negotiation needs take place: the sentence in question may have no weight in the fulfillment of each party's preferred choices. In contrast, when two parties negotiate, no debate has to take place: both parties can agree on all the relevant facts, yet find themselves in a situation where their choices are incompatible (indeed, zero-sum perfect information games can be seen as scenarios of this kind).

III.

If we cannot talk about normativity in only one sense, then we cannot talk about normative disagreement without qualifications. Thus, given the difference between Norms and norms*, between Normativity on the one hand and normativity* on the other, two basic kinds of normative disagreement seem to appear in the horizon. For example, imagine a discussion between two people, each of which defends one of the following:

C: the blockade against Cuba should be ended.
D: the blockade against Cuba should not be ended.

Prima facie, it appears as if C and D could not both be true. Moreover, it appears as if it would be inconsistent for someone to believe in both. Yet, in a world without normative facts, relativism renders both these

options possible. For example, imagine the following understanding of these sentences:

C*: Given facts [x,y,z...], and policy δ, the blockade against Cuba should* be ended
D*: Given facts [x,y,z...], and policy ε, the blockade against Cuba should* not be ended

Once a relativized sentence is presented explicitly it becomes a factual sentence to which a truth-value can be assigned. Since norms* are policies, C* will be true iff given the set of facts, policy δ dictates for the blockade to be lifted (ditto for D* after the relevant changes). Thus, I could agree with D*, in fact I could assert D*, even if I also believe and assert C*. A belief in either of them would then be a pure belief. Unlike cases of contextual disagreement, in which sentences are considered to be in disagreement iff they cannot both be true after being desensitized, C and D are in disagreement even in the case aforementioned.

Let us call a scenario such as the above an instance of normative disagreement*. Now let us follow this scenario from the first-person point of view. Your interlocutor has asserted D, and you have asserted C in reply. She asks you to explain yourself, to which you reply by asserting C*. She did not know you meant C*. All along she believed that you meant the opposite of what *she* had in mind when asserting D (that is, she thought you meant to deny D* by asserting that e.g. given facts [x,y,z...], and policy ε, the blockade against Cuba should* be ended)—and she tells you so. What kind of insight have you gained about the dynamic of this discussion after her announcement? It seems that you have learned that the two of you were not talking about the same thing—the two of you were talking past each other. Indeed, had it been a heated discussion, one of both of you would probably apologize for having pressed on with the debate. You may perhaps laugh a little at the end, thinking about how silly the whole scenario would have seemed to an observer who *knew* from the beginning what you have learned via her announcement. Looking back, it appears that there was *nothing about which* the two of you were disagreeing—even if your policies were in conflict. This is clearly an instance of equivocation.

In simple terms, speakers equivocate when, while presumably talking about the same thing, they are unknowingly talking about different things. Equivocation does not entail either agreement or disagreement, but it does convey the illusion of being in one of them. If two people were to get into a heated discussion over whether Harry's last presentation was on a given topic, just to finally realized that they were talking about two different people named "Harry", they would acknowledge they were not disagreeing—indeed, they were hardly communicating in a meaningful way (ditto were they to have been asserting their agreement with Harry's conclusions).

What is missing in the above case of conflict is something that we can find in instances of what we shall call "Normative Disagreement": agents R and S are in Normative Disagreement iff R holds a Norm α to be a fact, S holds Norm β to be a fact, and only one of [α,β] can be a fact. Of course, people can talk past each other in *apparent* instances of Normative Disagreement—which is to say that, for one reason or another, people can equivocate even when they think they are in this scenario. However, unlike normative disagreement*, at least some cases of what appears to be Normative Disagreement are such that, once each interlocutor has asserted her claim explicitly, both parties would agree they are debating over a *disagreement*. If you claim that C, and your claim is based on what you believe to be a Norm, and your interlocutor recognizes that what you think she says is indeed what she is saying, then it is not possible for you to accept your interlocutor's claim to the opposite if you both are in Normative Disagreement. For in Normative Disagreement your interlocutor would have to be asserting either that not-C or a claim that entails the falsity of C.

On the other hand, were your interlocutor's claim to be based on a norm* while your claim appeals to a Norm, then, upon having realized this piece of information, you would also realize that the two of you are not debating over a *disagreement*. In fact, given that your claim is a Normative claim while hers is a normative* claim, there needs not even be a conflict of preferences in this scenario—your assertion, and your Normative beliefs, needs not even be in line with your desires. Ultimately, the two of you are simply talking past each other. For what sort of debate can happen when only one person is asserting a sentence and what kind of conflict can there be

when there is only one choice being pressed forward? As the saying goes, it takes two to tango.

Your interlocutor could now argue against your appeal to Norms, and in doing so a debate would ensue on whether there is such a Norm, or any Norm for that matter. This would not be a normative disagreement of any kind, but a regular factual disagreement: your belief that there are Norms, and her belief that there are no Norms. Yet, this is also a substantial factual disagreement. Whether there are Norms or not is a significant fact—one that would dictate whether any Normative beliefs are true and, with it, whether there are any instances of Normative Disagreement in which one of the sentences in question is true.

But your interlocutor needs not be so philosophically oriented. Perhaps she refuses to engage in a discussion on normative realism while at the same time asserting that the two of you were in some sort of normative disagreement. Yet, in what way, if any, could she be right? The answer is none. *Ex hypothesi*, she is not claiming the two of you were in Normative Disagreement. If she is claiming that the two of you were in normative disagreement*, then she has not been engaging what *you* said—for your sentences appealed to Norms. And if she says that she *interpreted* your sentence as a normative* claim, then she would be corroborating your suspicions: she was not arguing against *your* claim. Of course she was free to interpret your claim as a normative* one—she was just mistaken in doing so. While she was talking about what her desires demand, you were asserting what you believe the world commands.

IV.

The world-oriented phenomenology described above can be seen as indicative of an impulse that permeates throughout the entire practice of epistemology. Our desire to understand how we know rests on a particular notion of what there is to be known. Non-realist conceptions of normativity were attacked above for their inefficient account of this very same underlying notion. Yet, nowhere is this tacit demand (namely, the demand for a

metaphysically significant picture of the world to be the bases of our epistemology) better seen than in the debate on global skepticism. In this section I will argue that the true object of global skepticism is the metaphysical presumption of an intelligible world.

It is commonly argued that global skeptical arguments are self-defeating. For the purpose of this essay it is sufficient to define the notion of a defeater in the following terms: An argument (or proposition) y is a defeater of another argument (or proposition) x if and only if were y to obtain then x could not obtain. The term "obtain" is here being used to indicate either soundness (in the case of an argument) or a truth-value of true (in the case of a proposition).

Global skepticism is the thesis that, for reasons involving the need to meet justificatory desiderata, we are not justified to claim knowledge. This thesis was most famously propounded by the Pyrrhonian skeptics of the Hellenistic period. Global skepticism has been accused of being self-defeating. Indeed, it appears that the Pyrrhonians themselves gladly welcomed the apparent inconsistency of their position. In his Outlines of Skepticism (Empiricus 2000), Sextus Empiricus presents us with a global skeptical argument to rule them all: the proof against proofs. This argument tries to show that, if the argument is sound, then there are no proofs. Since, presumably, the argument itself is a proof in case of being sound, it seems to follow that we have a proof that there are no proofs (Empiricus 2000, pp. 2. 144-184). However, it is not necessary to dwell on the casuistry of particular skeptical arguments in order to discuss their underlying features—it is sufficient for our discussion to recognize the general structure of such arguments and its apparently self-defeating features. Global skeptical arguments such as the proof against proofs seem to have the following form:

Schema of the Proof against Proofs (SPP):

1) Every argument has at least one defeater argument.

2) In order for us to be justified to believe in the soundness of any argument over that of its defeaters we must meet some *justificatory desiderata*

3) We cannot meet the *justificatory desiderata*.

4) Therefore: we are not justified to hold any argument to be sound.

Given our criterion of defeat, for an argument y to count as a defeater of SPP it must be the case that, if y obtains, then SPP cannot obtain—that is, if y is sound then SSP cannot be sound. This renders it evident that SPP is not a defeater of SSP. Neither does SPP's conclusion deny any of its premises, nor is such a negation implied by it. Moreover, SPP's conclusion does not negate itself, nor does it imply such a negation (ditto for SPP's premises).

However, there is an intuitive uneasiness in thinking of arguments like SPP as informative or, more disturbingly, as tenable. That global skeptical arguments such as SPP are not self-defeating in the strictest sense of the term (as least as defined above) does not mean that they can be held—the very nature of SPP prevents any epistemic agent from assenting to it. Let us assume the soundness of SSP. If it is the case that justification is necessary for a true belief to count as knowledge, then since we cannot be justified in believing any argument to be sound (or any thesis to be true) we cannot claim to know of any argument that it is sound (or of any thesis that it is true). Thus, neither can we claim to know any argument to be sound nor can we claim to know any thesis to be true. Yet, this means that we cannot claim to know that SSP is sound. This effectively means that, if SSP is sound, then we cannot know that SSP is sound—in other words, if SSP is sound, then its soundness is unknowable.

Those who accuse global skeptical arguments such as SSP of being self-defeating may be better understood as portraying such defeat within the framework of epistemic goals and pragmatic grounds. If the goal of inquiry is the acquisition of knowledge, then arguments such as SSP cannot be a tool for the fulfillment of our epistemic ends for it is not possible to hold an argument such as SSP in a consistent fashion. If we wish to hold only justified beliefs and, having assessed all the relevant evidence, we come to the conclusion that we are justified to believe in

the soundness of SSP, then we have effectively overrun our own grounds for believing in the soundness of SSP. We cannot continue to uphold SSP to be sound, for we now acknowledge that we cannot be justified to believe SSP to obtain in the place of any of its defeaters. Thus, if we are justified to believe SSP to be sound, then we are justified not to believe SSP to be sound. Attempting to uphold global skepticism via the assertion of arguments such as SSP would be the epistemic equivalent of attempting to utter "I am not uttering a sentence right now".[2]

Notice that the above is in no way inconsistent with SSP. Both arguments show that, if justification is necessary for knowledge and either one of these arguments is sound, then we could not come to know that SSP is sound were it to be sound. However, this agreement can hardly be seen as indicative of an agreement on the lessons to be derived from the scenario. Deriving a contradiction from a set of premises has long been seen as an efficient sorting strategy when it comes to accepting or rejecting beliefs. The impossibility of holding arguments such as SSP in a consistent fashion seems to dictate for the rejection of something—yet this something cannot be the possible soundness of SSP. In rejecting the claim that we can hold SSP consistently we are responding to the apparent unintelligibility of a world where SSP is sound.

By assuming the soundness of SSP we are forced to contemplate the possibility of an unknowable world. This scenario becomes even more frightening once we realize that the self-referential features of global skepticism deprive us of reasons to justify this epistemically nightmare-esche scenario itself. Yet, and in a truly Damoclean fashion, we can acknowledge that the possibility of an argument such as SSP being sound is a constant , if passively latent, threat to our dearest epistemic goals. If, in presenting an argument such as SSP, it is the desire of an epistemic agent to allow her interlocutor to experience what it is like to reflect on such an incomprehensible scenario, we cannot expect to hold the former's success as an objection to his argument.

[2]For more on pragmatic self-refutation see (O'Connor 1951). Sextus Empiricus' pyrrhonian skepticism tends to be seen as prey of pragmatic self-refutation. McPherran provides a list of sources on the topic as well as an illuminating analysis of the issue; see (McPherran 1987).

If reductiones ad absurdum are useful because they allow us to sort the wheat from the chaff, then rejecting SSP on the grounds that the assertion of its soundness produces an insurmountable inconsistency could be seen by some as a legitimate move. However, as it was just shown, this strategy is not equivalent to refuting the soundness of arguments such as SSP. An attempt to deny that SSP is sound cannot be based on that, were it to be sound, its soundness would entail our incapacity to be the kind of epistemic agents we *think* we are. If a historical analogy is adequate we could then portray global skeptical arguments as standing closer to Zeno of Elea's elaboration of Parmenides' arguments than to the apparently self-defeating global relativism of thinkers like Protagoras. That such a comparison is relevant can be seen once we inquire upon the alleged self-refuting characteristic of global relativism: An assertion to the truth of all theses entails the acceptance of that very same thesis' negation as true.[3] Thus, global relativism is self-defeating in accordance to our definition of the term. In contrast, Parmenides' and Zeno's arguments seek to convey the picture of a less-than-immediately-intelligible world. The counterintuitive nature of their results stands in direct opposition to the rawest, and most immediate, notion of the world shared by us all without therefore falling in self-defeat. In doing so, they do not show that we do not conceive of the world in a certain way. Instead, they appeal to our picture of the world in order to disprove its feasibility. A mere reassertion of our experiences cannot be seen as a feasible strategy when trying to disprove the soundness of, for example, Zeno's argument against motion. [4]

In this sense, those puzzling features of global skepticism which prevent us from being capable of holding the position consistently separate it from other apparently untenable positions like, for example, the normative non-realist described earlier in this essay. Whereas the failure of the normative non-realist positions emerges from its inability to explicate, and account for, epistemic phenomena (e.g. disagreement),

[3] For the best-known version of this argument see Plato's *Theaethetus*, 171a-c (Plato 1997). For a more contemporary as well as inquisitive discussion on Protagoras' relativism see Mi-Kyoung Lee's *Epistemology After Protagoras: Responses to Relativism in Plato, Aristotle, and Democritus* (Lee 2005).

[4] For more on Parmenides' view see (Curd and McKirahan 2011).

the global skeptic purposely aims at placing us in an apparently unintelligible position. Indeed, the relevance of such a move can be understood only in contrast to the most general, and basic, objections that can be raised in the dialectic on global skepticism. An objector could argue that, on the one hand, the mere logical possibility that SSP may obtain provides little to no grounds for judging it worthy of inquiry and discussion. On the other hand, given the repercussions of SSP's soundness, there seems to be no reason to discuss something which appears to dictate for quietism.

However, these latter arguments represent a radical shift in the scope and nature of the discussion. The issue is no longer whether global skeptical arguments are self-defeating; what is now requested is a stance on whether the consequences of global skeptical arguments are desirable, let alone endurable. [5] A judgment on this topic is a judgment on how much weight, if any, we are to place upon the possibility of there being truths which are necessarily unknowable; it is a judgment on the degree to which we wish to allow such possibility to influence our notion of inquiry[6]—it is a judgement over our most basic conception of the world. By presenting us with a scenario where our unrestrained thirst for knowledge may prove the world to be a dry riverbed, global skeptical arguments pin our desire for truth against the ropes—and even those who allegedly claim to love deserted landscapes retreat from such arid a world. The unknowable truth that would follow from these arguments is one which demands for a reassessment of our own self-image as inquirers who conceive the object of their epistemic desires to be within reach. In judging upon the relevance of this topic we are passing judgment upon our very own notion of inquiry—we are passing judgement over our most elemental conception of the world.

If anything can be seen as a default epistemic stance is, *pace* Protagoras, a rejection of complete infallibility. Not only do we accept the possibility of some, if not all, of our beliefs being false, but such possibility is also conceived as nothing more than a mere bump in the road to knowledge. Yet, the gap between the possibility of being wrong

[5] This statement is not to be confused with the anti-skeptical objection concerning the impossibility of living a skeptical life (for a classic discussion on the topic see section XIII of Hume's *An Enquiry Concerning Human Understanding* (Hume and Millican 2007).
[6] For an example of this kind of approach see (Rescher 2007, 39-50).

and the impossibility of being right is one which can hardly be bridged by appeals to fallibilism and pragmatic considerations. If our notion of error depends on our notion of "getting things right" in a realist-like sense of the term[7], if our notion of objectivity is indeed based on a presupposition of realism as the beginning of our discussions[8], then the impossibility of having justified belief would render the world unintelligible in any meaningful sense. An assertion to the opposite, a reiteration of our conviction in the intelligible nature of the world and experience, can hardly be expected to eliminate the fear caused by a possible state of necessary ignorance.[9] It may provide a reason to search for alternative frameworks within which to discuss the topic, but it will not do away with it. Assuming the intelligibility of the world when it comes to choosing a path away from global skeptical arguments provides little guidance *other than* a proviso for this path to be paved by intelligibility—even if as nothing more than an implicit desideratum. This requirement, however, is insufficient, for it can be fulfilled with ease by incompatible goals, methods, and even answers within the same conceptual framework.[10]

It is upon a canvas of a strong conception of truth that the notions of error and objectivity, as described all throughout this essay, colour the landscape of an unintelligible world. This is why alternatives to a strong

---

[7] (Rescher 2007, 27-34; 80-84).

[8] See Barry Stroud's discussion of the topic in chapter II of his book *The Significance of Philosophical Scepticism* (Stroud 1984).

[9] That a conviction is not sufficient to dispel doubt is precisely what is at issue in certain takes on Moore's argument against skepticism about the external world. See, for example, Unger's argument against the latter in his book *Ignorance* (Unger 1975, 20-36). Unger's argument relies on our intuitive response to counterfactual scenarios in which we are shown to be mistaken in having claimed certainty. We may conceive of Unger's argument as appealing to the phenomenology of the situation: in Unger's discussion, that we may come to experience embarrassment upon discovering we were wrong is a consequence of recognizing our conviction that we are entitled to claim certainty to be misguided. Embarrassment is not eliminated by our conviction, it is caused by it. Similarly, the conviction that the world is intelligible does not dispel the puzzlement that characterizes our apprehension of the possibility that SSP may be sound. For a defense of Moore's argument which emphasizes its appeal to our notion of plausibility and certainty see (Lycan 2001).

[10] For historical examples on the employment by both sides of a debate of global-like skeptical arguments, as well as on the apparent undesirable consequences of holding skeptical arguments, see (Popkin 1964, 1-16; 197-217).

notion of truth can be seen as responses to global skeptical arguments[11]. In attempting to preserve some resemblance of intelligibility in the world, these options provide us with an alternative within which we may conceive of inquiry as conducive to knowledge rather than paradox. However, this move is not required in order to defend the intelligibility of a knower-friendly world from global skeptical arguments. Inquiry can be portrayed as indeed conducive to an unintelligible world—reason for which inquiry itself may be seen as an inadequate tool for the apprehension of truth.[12]

Conclusion

It would be a gross misinterpretation of this essay to conceive of it as a study on ordinary language. The terms "conflict" and "disagreement" could be replaced for each other throughout the essay *salva veritate*. What matters to the argument is not whether we should call the phenomenon to which we refer by the term "disagreement" or "conflict". What matters is that these phenomena *are* different—to the point where a reduction of one to the other would be not only impractical, but also misguided. Those who wish to reduce normative disagreement to a conflict between policies need to explain why we should dismiss the phenomenological evidence we have that the non-realist would be misinterpreting what we say. Moreover, the notions of instrumental and non-instrumental normativity need not be mutually exclusive. The kind of value expressed by each is consistent with that of the other. The problem is not with the notion of instrumental normativity, the problem is with the claim both that there is only instrumental normativity and that we can accommodate the phenomenology of discourse within a framework characterized solely by instrumental normativity. This is the claim to which I object. For the phenomenology of engaging in normative discourse from the Normative

[11] Rorty's notion of truth as a commendatory term can be seen as another example of an alternative understanding of truth which aims at overcoming some of the problems described above (see (Rorty, Solidarity or Objectivity 1985)). For a better understanding of Rorty's view on the applications of the notion of truth and its consequences see also (Rorty 1988).

[12] This is precisely the kind of move perform by fideists. See (Popkin 1964, 34-38; 92-99).

stance can accommodate that of engaging in normative discourse from the normative* stance, but not *vice versa*. The non-realist needs to show unequivocally why we should not take seriously these serious cases of equivocation. This, from the first-person point of view, she has done not.

Moreover, this phenomenology is indicative of a much deeper attitude towards the world—an attitude the constitutive role of which is demonstrated by global skepticism. The somewhat disturbing, almost incomprehensible nature of the global skeptical position emerges not from any sort of content-based inconsistency—that is, it is not due to explicit or implicit contradictions in its claims. Global skepticism elicits such reaction via the manner in which it attacks our presupposition of an intelligible world—presupposition which is indeed constitutive of the epistemic enterprise. Yet, this presupposition is metaphysical in nature, and it demands for a world that is constituted in such a way that we must be able to comprehend it. It is a metaphysical demand for an intelligible world; that the content of such demand can be questioned in a consistent fashion in epistemology ought to be seen as a meaningful sign of the connection between the latter and metaphysics. Those who dream of an epistemology without metaphysics dream of square epistemic circles. Whatever value we may attach to epistemology must be indicative of the value we must attach to metaphysics.[13]

Literature

Curd, Patricia, and Richard D. McKirahan. *A Presocratics reader : selected fragments and testimonia.* 2nd edition. Indianapolis: Hackett Pub. Co, 2011.

Empiricus, Sextus. *Outlines of Scepticism.* Edited by Julia Annas and Jonathan Barnes. Cambridge / New York: Cambridge University Press, 2000.

Field, Hartry. "Epistemology Without Metaphysics." *Philosophical Studies* 143 (2009): 249-290.

Gibbard, Allan. "Rational Credence and the Value of Truth." *Oxford Studies in Epistemology 3*, 2008 .

———. *Thinking How to Live.* Cambridge, MA: Harvard University, 2003.

———. *Wise Choices, Apt Feelings.* Cambridge, MA: Harvard University, 1990.

[13] I wish to thank Emily Hodges for her invaluable help in the process or writing this essay—help that manifested in, but cannot be reduced to, conversation, email, and the proofreading of drafts. Needless to say, any mistakes in this essay are solely of my authorship.

Hume, David, and P. F. Millican. *An enquiry concerning human understanding / David Hume ; edited with an introduction and notes by Peter Millican.* Oxford / New York: Oxford University Press, 2007.

Lee, Mi-Kyoung. *Epistemology after Protagoras: responses to relativism in Plato, Aristotle, and Democritus.* Oxford / New York: Clarendon Press , 2005.

Lycan, William G. "Moore against the New Skeptics." *Philosophical Studies* 103, no. 1 (2001): 35-53 .

MacFarlane, John. "Making sense of relative truth." *Proceedings of the Aristotelian Society* 105 (2005): 321– 339.

———. "Nonindexical contextualism." *Synthese* 166, no. 2 (2009): 231-250.

McPherran, Mark. "Skeptical Homeopathy and Self-refutation." *Phronesis* 32, no. 3 (1987): 290-328.

O'Connor, D.J. "Pragmatic Paradoxes and Fugitive Propositions." *Mind* LX, no. 240 (1951): 536-538.

Plato. "Theaethetus." In *Complete Works*, edited by John M. Cooper and D.S. Hutchinson, translated by M.J. Levett, 157-234. Indianapolis, IN: Hackett Pub, 1997.

Popkin, Richard Henry. *The history of scepticism from Erasmus to Descartes.* New York: Harper, Torchbooks, 1964.

Rescher, Nicholas. *Error.* Pittsburgh, PA: University of Pittsburgh Press, 2007.

Rorty, Richard. "Representation, Social Practise, and Truth." *Philosophical Studies* 52, no. 2 (September 1988): 215-228 .

Rorty, Richard. "Solidarity or Objectivity." In *Post-analytic philosophy* , by John Rajchman and Cornel West, 3-19. New York: Columbia University Press, 1985.

Stroud, Barry. *Engagement and Metaphysical Dissatisfaction.* Oxford / New York: Oxford University Press, 2011.

———. *The significance of philosophical scepticism.* New York: Oxford University Press, Clarendon Press, 1984.

Unger, Peter K. *Ignorance: a case for scepticism .* Oxford: Clarendon Press, 1975.

On the Prospects for Naturalism

Nicholas Tebben
(Johns Hopkins University)

Contemporary naturalism has two components. The first is ontological, and says, roughly, that all and only what the sciences say exists, really does exist. The other is methodological, and it says that only scientific explanations are legitimate explanations. Together these commitments promise a coherent picture of the world that is nicely integrated with an attractive epistemology. Despite the obvious appeal of naturalism, I would like to sound a note of caution. First, I would like to argue that naturalism's ontological commitment cannot be vindicated. Not, that is, that it is false; rather, I argue that any attempt to *show* that it is true presupposes that it is not. Second, I argue that methodological naturalism is false. But, again, the problem is not straightforward. I will not claim that there are gaps in the explanations offered by science, such that the scientific project would be incomplete without emendation. Instead I argue that the goodness of an explanation depends, in part, on how the event to be explained is described, and that, some events, under some descriptions, call for non-scientific explanations.

1. Introduction

Contemporary metaphysics, in what is perhaps its most popular form, finds its home in a larger set of views that promises a unified picture of the world and our place in it. Collectively known as 'naturalism', this set of views proceeds from the conviction that empirical science is our only guide to the nature and content of the world. Naturalism has two main components. The first is ontological[1]:

[1] Despite its presentation here, the precise content of naturalism is controversial. The thesis that I call 'ontological naturalism' is a thesis of intermediate strength, relative to the other "naturalistic" metaphysical theories in the field. Physicalism, for example, is a stronger thesis. As far as I can see, however, none of the plausible ways in which it might be modified would bear on the argument to be developed below.

Ontological Naturalism: The kinds of things that exist are all and only those kinds that play a role in our best sciences, or will play a role in successor sciences that are very similar to those currently in the field.[2]

The second component of naturalism is a thesis about what constitutes a good explanation:

Methodological Naturalism: All legitimate explanations are scientific explanations.

Together the two claims present a coherent whole. We believe in things for explanatory purposes, and if all legitimate explanatory purposes are scientific ones, then all things for which one has legitimate grounds to believe are those necessary for science. If these two commitments are correct, then philosophy and all other branches of inquiry are, to the extent that they are legitimate, branches of the empirical sciences.

Naturalistic philosophy usually proceeds by finding some apparently problematic, but philosophically significant, subject matter, and providing an account of the nature or functioning of the subject matter in question that appeals only to members of ontological categories that find employment in the sciences. Ultimately, if the philosophical project is successful, the explanations that this account will provide will be scientific explanations. In completing such a project, the naturalistic philosopher helps develop a coherent picture of the world, integrated by scientific principles and commitments.

---

[2] It is worth noting that this is a *distinctively* metaphysical thesis, and not a scientific one. Here is Stroud, on this point: "A positive science of physical nature tells us what the world is like – what qualities objects in the world do, in fact, have. But that physical science alone does not establish [that only physical objects exist]. Physical scientists professionally restrict their attention to the physical aspects of the world that can be captured in their theoretical network. If there is more to the world than that, physical science says nothing about it. The metaphysical theory of atomism or physicalism goes one step further. It says that atoms or the physical qualities of things are all there is." (Barry Stroud, *The Quest for Reality: Subjectivism & The Metaphysics of Color* (Oxford: Oxford University Press, 2000), 10.) Moreover, it is not an accident that the sciences remain uncommitted about those things that fall outside of their theoretical network. If there were any such things, there could be, by their very nature, no scientific evidence that would bear on the hypothesis that they exist.

Despite its obvious attraction, I am pessimistic about the prospects for naturalism. There are two reasons for my pessimism. First, I will argue that there are reasons to believe that a demonstration of the ontological thesis cannot, even in principle, be completed. This is not to say that Ontological Naturalism is false; I do not argue that there *is* something that does not find a place in the scientific picture of the world. Rather, I will argue, the problem for naturalism on the ontological side is that any *demonstration* that it is true requires that one presuppose that it is not.[3]

The other challenge that naturalism faces comes from the methodological direction. The problem, in short, is that explanatory contexts are referentially opaque. So even if there is a scientific explanation of every fact and event, it does not follow that there is a scientific explanation of every fact and event *under every description*. Indeed, I will suggest that there are some facts, under some descriptions, that call for explanation in terms of categories that cannot be shown to be identical to those familiar from the sciences.[4] If that is so, then the appeal of naturalistic metaphysics will be much weakened, and the plausibility of the picture of inquiry as unified by scientific procedures and explanations much reduced.

---

[3] There is a clear affinity between this argument and some of Putnam's arguments. He has, for example, long argued that scientific practice essentially involves appeal to epistemic values, such as coherence and simplicity. He takes this fact as a premise in an argument that there is no sharp dichotomy between facts and values. See Hilary Putnam, "Beyond the Fact/Value Dichotomy," in *Realism with a Human Face*, ed. James Conant (Cambridge: Harvard University Press, 1990), especially pp. 138-141, and Hilary Putnam, *The Collapse of the Fact/Value Dichotomy, and Other Essays*, (Cambridge: Harvard University Press, 2002), especially chapter 2. I think that the point about scientific practice generalizes: providing *any* argument presupposes that there are, what Putnam would call, "epistemic values". I am, however, less ambitious than Putnam. My contention is only that this fact will allow us to show that Ontological Naturalism cannot be vindicated.

[4] Again, this is *not* to say that they are not identical to such categories. They may well be. But they cannot be *shown* to be identical to naturalistic categories.

2. Naturalizing the Epistemic

Normative facts and properties are the most notable among those that resist integration into a naturalistic world-view. If Ontological Naturalism is to be vindicated, then, either normative facts and properties must be shown, despite initial appearances, to be identical to natural facts and properties, or else it must be shown that they do not exist at all.

Whatever its merits as a means of dealing with other normative kinds,[5] eliminativism about epistemic normativity seems to be ill advised.[6] The problem, in brief, is that any *argument* for epistemic eliminativism would need to provide good reasons to believe that its conclusion is true, but epistemic eliminativism just is, in part, the claim that there are no such things as *good reasons to believe*.[7]

More promising than attempting to eliminate the epistemic, however, are attempts to reduce epistemic facts and properties to facts and properties that are employed by the sciences. And, in fact, there have been many attempts to reduce one or another kind of epistemic fact or property. In isolated cases, I do not doubt that this can be done. Vindicating Ontological Naturalism, however, requires not that *some* epistemic facts and properties be reduced to natural ones, but that *all* of them are so reduced. It is this project, that of reducing all epistemic categories, that cannot be completed. And this project cannot be completed not for contingent reasons regarding, for example, its difficulty, or our lack of imagination. I will argue that, even in principle, a wholesale reduction of the epistemic to the natural is not possible.

Though reducing *A*'s to *B*'s involves showing that *A*'s and *B*'s are identical, reductions are typically not reductions without loss. Reducing one thing to another involves re-conceptualizing at least one of the involved facts, properties or objects. In particular, reductions typically

[5] For a discussion of its merits in the moral case, see J.L. Mackie, *Ethics: Inventing Right and Wrong* (New York: Penguin Books, 1977).

[6] And, indeed, epistemic eliminativism has no defenders.

[7] Notice that it would not do simply to note that the fact that there are no epistemic facts is entailed by some other fact *F*, and that *F* is true. For if there are no epistemic facts, then the fact that both a conditional and its antecedent are true provides one with no *reason* to conclude that its consequent is true. Without any epistemic facts, the inference is optional.

involve denying that the kind being reduced has features other than those that are characteristic of the kind to which it is being reduced. For example, a physical reduction of the mind involves not simply saying that the mind and (some favored part of) the body are identical, but that minds are not immaterial. If a reduction was *simply* a matter of finding that the two kinds are identical, the possibility would be open that a successful reduction of the mental to the physical would show that there are immaterial physical objects. But of course nothing of the sort is on offer. A reduction of the mental to the physical crucially requires saying that the mental does not have all of the properties that it was previously thought to have.

A similar impoverishment is characteristic of attempts to provide a naturalistic reduction of epistemic normativity. As we ordinarily conceive of them, epistemic categories carry deontological significance. One violates a *duty* when one believes without justification; in doing so one behaves *irresponsibly*. In virtue of such facts, epistemic evaluations have an imperatival aspect. To say that one would not be justified in believing that *p* is, in part, to *forbid* one from believing that *p*. Likewise, holding a justified belief is a matter of *acting in accord with one's duty*. And saying that one would be justified in believing that *p* is, in part, to *permit* one to believe that *p*.

But because the natural world contains no deontic facts, naturalistic reductions of epistemic categories are reductions to something that is non-deontic. And in completing a reduction of an epistemic category to a natural category, one thereby finds that the epistemic category in question does not, in fact, possess the deontic characteristics it is usually thought to possess. The most popular, and promising, naturalistic reductions of epistemic categories are the various reliabilist epistemic theories.[8] Though the details are many and varied, the heart of the view

[8] Kornblith, who offers a reliabilist account of knowledge, denies that his view is reductionistic: "Epistemic terminology, and, indeed, philosophical terminology in general, must be grounded in the world if it is to be naturalistically legitimate. This does not require that such terminology appear in our physical theories, for naturalists need not accept any sort of reductionism. ... Naturalism would only threaten to eliminate epistemic terminology as illegitimate if there were no prospect of discovering theoretically unified epistemic phenomena." (Hilary Kornblith *Knowledge and Its Place in Nature* (Oxford: Oxford University Press, 2002), 25.) The kind of "reductionism" that Kornblith has in mind is a particularly strong variety, one according to which all legitimate explanation is

is that epistemic properties – like *being justified in believing* – are identical to natural properties concerning the reliability (under some set of conditions or other) of one's belief-forming processes. Such views are naturalistic precisely because reliability, understood as a high ratio of true to false beliefs, is no more problematic than belief and truth, both of which naturalistic philosophers have made peace with. These notions, which have a place in a broadly scientific understanding of the world, contrast with that of justification, understood as conformity to one's duty, which does not. So if it has been (or will be) found that, say, *being justified in believing*, is the same property as *being the product of a reliable belief-forming process* (under some set of conditions), it has been, or will be, found that epistemic normativity is radically different than was previously thought.

The difference is even more apparent in the case of *un*justified beliefs. Assume that a belief is unjustified just in case it is not justified,[9] in that case, unjustified beliefs are those that are formed through belief-forming processes that are either not reliable, or not used in the right conditions. If to be unjustified is to be the product of an unreliable process (or a reliable process used in the wrong circumstances), then the imperatival aspect that we find in epistemic evaluations is misplaced. For from the fact that a belief is unlikely to be true, it does not follow that it one is prohibited from holding it, though from the fact that holding the belief requires that one violate one's duty it *does* follow that holding that belief is prohibited. Naturalistic epistemology is a revisionary project.

This is important because vindicating Ontological Naturalism requires an *argument*, one which makes it *rationally mandatory* to hold that the only things that exist are those things that play a role in our best current sciences, or close successors to those sciences. But the natural world contains no facts about what is mandatory. And so it is precisely this aspect of epistemic facts – that they do, or can, concern what is rationally mandatory – that is not preserved in a naturalistic reduction.

---

not merely scientific explanation, but explanation at the level of physics. Kornblith is right that naturalists need not accept this extreme view. But they *must* be reductionists in some sense, at least in that they must see (as, indeed, Kornblith does see) epistemic categories as identical to the categories, or at least constructions thereof, recognized by the sciences.

[9] That is, assume that there are no beliefs that are neither justified nor unjustified.

That it is not preserved is what allows us to say that a reduction of the epistemic to the natural really is a reduction *of* the epistemic *to* the natural: if the reduction is successful, those aspects of epistemic facts that do not fit nicely into a naturalistic understanding of the world will be discarded.

But this fact is also the reason that Ontological Naturalism cannot be vindicated. If an argument in favor of Ontological Naturalism were successful, it would effect a reduction of the epistemic to the natural. But if that reduction was to go through, it would show that epistemic facts cannot be facts about what is rationally mandatory, as natural facts have no imperatival aspect. They must *merely* be facts about, for example, which beliefs were produced by reliable belief-forming processes. But if the conclusion of the reductive argument is, in part, that epistemic facts are *merely* facts about, say, the reliability of belief-forming processes, then the premises of the argument cannot make its conclusion rationally mandatory. And so, *by its own lights*, we are free to accept the premises of the argument for Ontological Naturalism, without accepting its conclusion.

It is worth emphasizing that the naturalist program requires more than showing that there is a sound argument with Ontological Naturalism as its conclusion. The mere fact that an argument is sound does not, in general, make its conclusion rationally mandatory. Indeed, a sound argument need not make its conclusion rationally mandatory, even if one *knows* that it is sound. Consider: it is necessarily true that the set of real numbers is uncountable. Since it is necessarily true, any argument for the proposition that the set of real numbers is uncountable is valid and, if its premises are true, sound. But few such arguments make the conclusion rationally mandatory. Indeed, imagine that I assure you that the set of real numbers is uncountable. You then *know* that the argument "grass is green, so the set of real numbers is uncountable" is sound. Nevertheless, it is not the argument that gives you conclusive reason to believe that the set of real numbers is uncountable; your reasons, rather, comes from my testimony. Demonstrating that Ontological Naturalism is true requires a sound argument, which reduces the epistemic to the natural, but which *also* makes believing the conclusion rationally non-optional, given that one believes the premises. And this is what the naturalist cannot have, as

admitting that there are natural properties that can make holding some belief *mandatory*, or such that holding that belief is the only way to act in conformity with one's *duty*, would make a mockery of the idea that the epistemic had been *reduced to* the natural.

This fact is obscured in the writings of those working in naturalized epistemology, because the problem would arise only in the attempt to reduce *all* epistemic categories to natural categories. Those who actually attempt a reduction do so one fact or property at a time, and so they are free to make use of unreduced epistemic facts or properties in their reductions. Consider Kornblith's reduction of knowledge to reliably produced true belief. He begins by using pre-theoretic intuitions to roughly specify a property of interest, and then looks to see if empirical inquiry finds anything that can be recognized as relevantly related to the property that was identified pre-theoretically. As empirical inquiry proceeds, its needs take priority over the dictates of pre-theoretic intuition. Kornblith argues that empirical inquiry does, indeed, identify something recognizable as relevantly related to the property pre-theoretically identified as knowledge. In particular, cognitive ethology utilizes something recognizable as knowledge ascriptions in the explanation of animal behavior. Moreover, the knowledge that cognitive ethologists attribute to the animals that they study can be recognized as a natural kind through the role that it plays in fixing traits through natural selection. Since it is a natural kind, it can be investigated inductively. Ordinary empirical inquiry then proceeds, and in the course of this inquiry one discovers that the property cognitive ethologists need for their evolutionary explanations of trait fixation is merely that of a reliably produced true belief. Hence, Kornblith concludes, knowledge is reliably produced true belief.[10]

Kornblith may well be right about this, and his argument itself may also be successful. But the reason that he does not encounter the problem that I have identified is that he attempts to reduce only one epistemic category, and so he is free to make use of others. The problem arises only with the attempt to reduce all epistemic categories, or the epistemic *as such*, to the natural. Those working in naturalistic

[10] This is the central argument of Kornblith's *Knowledge and It's Place in Nature*. His methodology is developed and described on pages 11-63.

epistemology, like Kornblith, achieve the success that they do precisely because they attempt only one reduction at a time. But naturalism is compelling because it promises a unified picture of the world, and in so far as naturalistic epistemology is important, it is so because it contributes to this unified picture.

What I have argued is that this picture cannot be completed. It cannot be completed all at once, nor can it be completed by conjoining local naturalistic projects, like those pursued by Kornblith, or Alvin Goldman. Again, a successful argument is not merely a sound one; it must be such that, if one accepts its premises, then one is *required* to accept its conclusion. So if Kornblith's argument is successful, it must be that *if* the premises are accepted, one is (epistemically) required to accept the conclusion. Otherwise, his audience could, without blame, accept the premises and reject the conclusion. But any reduction of the epistemic to the natural must relieve the epistemic of its, as it were, imperatival aspect. So Kornblith's reduction of knowledge to reliably produced true belief, presupposes that there is *some* unreduced epistemic property or other. And any argument that attempted to reduce *that* property would, if it was successful, presuppose that there is some *other* unreduced epistemic property. And so on. The problem is not with trying to naturalize epistemic facts or properties, it is with trying to naturalize *all* of them, either all at once, or in succession.

I should be clear about what this argument attempts to prove. Nothing that I have said indicates that there *are* any non-natural epistemic facts or properties. I have also, in the interest of not begging any questions, not claimed that one cannot know that Ontological Naturalism is true. If knowledge really is reliably produced true belief, and if one's belief that Ontological Naturalism is true is both true and reliably produced, then one knows that it is true.[11] I *have* argued, however, that there could not be a successful *argument* for Ontological Naturalism. A reduction of the epistemic to the natural would be an argument that purports to show that, if its premises are accepted, one is rationally required to accept that the facts and properties described in epistemic terms are identical to natural facts and properties. But if this is a reduction *of* the epistemic, *to* the natural, then it would show that

[11] I would like to thank Matthew McCauley for reminding me of this point.

epistemic facts and properties do not possess an, as it were, imperatival aspect. But if they do not, then the premises of the reductive argument do not rationally mandate that one accept the conclusion of the argument.

3. Explanation

Perhaps the argument to this point need not be a cause for much concern. Reducing the epistemic to the natural requires supposing that there is an irreducible epistemic property, but if that property plays no role in explaining any facts or events, then the naturalist's methodological commitment remains untouched. And that philosophy ought to be a branch of the sciences is, perhaps, naturalism's most fundamental commitment. So in closing, I would like to briefly discuss why the inability to complete the ontological component of the naturalist's project should undermine our confidence in its methodological component.

The heart of the case against Methodological Naturalism is that explanatory contexts are referentially opaque. Referentially opaque contexts are those in which substitution of co-referring expressions can change the truth value of the sentences in which the substitution occurs. The paradigmatic example of a referentially opaque context is the object position of a propositional attitude ascription, but explanatory contexts function in the same way. Putnam found a characteristically vivid way to illustrate this point:

> Suppose a professor is found stark-naked in a girl's dormitory room at midnight. His being naked in the room at midnight - ε, where ε is so small that he could neither get out of the room or put his cloths on between midnight - ε and midnight without traveling faster than the speed of light, would be a 'total cause' of his being naked in the girl's room at midnight; but no one would refer to this as the 'cause' of his presence in the room in that state. ... In its ordinary sense 'cause' can often be paraphrased by a locution involving *explain* ... The forest fire is *explained* (given background knowledge) by the campfire's not having been extinguished; but the professor's state at midnight - ε is not what we consider an *explanation* of the state of affairs at midnight.[12]

[12] Hilary Putnam, "Why There Isn't a Ready-Made World," in *Realism and Reason: Philosophical Papers Volume 3*, (Cambridge: Cambridge University Press, 1983), 213.

Now, there are many ways to describe the situation that Putnam imagines. A description of it, *given purely in the language of physics*, may well call for an explanation in terms of what could or could not happen in an interval of length ε. But a description of the situation given in terms of the professor and his clothes does not. Under such a description, a good explanation of the situation must make reference to, say, the professor's romantic involvement with a student.

The point generalizes. Describing a fact or event in one way rather than another makes some features of the fact or event contextually salient, and these contextually salient properties determine what kind of explanation is called for. A purely physical explanation of the professor's presence in the dorm room will, if the situation is described in those terms, be no good, because from a physical perspective, that the atoms under discussion *comprise* a professor and a dorm room is sheer accident. An explanation in terms of the professor's relationship with the student, however, is a good one, because it is essential to the explanation that the objects under discussion have their contextually salient properties.

This is important, because it means that *even if there is a scientific explanation for every fact and event*, nevertheless, there is room for non-scientific explanations of the very same facts and events, though under other descriptions. Above we saw that Ontological Naturalism cannot be validated, because *arguing* for it presupposes that there is some non-natural property, something like *being rationally mandatory*. I would now like to suggest that this property has a role to play in some explanations, of some phenomena, under some descriptions.

Stroud has long argued that epistemic externalism (like Kornblith's reliability theory) does not answer to our philosophical needs because it does not provide the kind of self-understanding that we seek.[13] The opacity of explanatory contexts allows us to explain Stroud's dissatisfaction. Assume that perception is reliable, and that the fact that human knowledge, in general, is possible, is identical to the fact that a creature with a certain set of cognitive faculties manages to accurately

[13] See Barry Stroud, "Understanding Human Knowledge in General," in *Understanding Human Knowledge* (Oxford: Oxford University Press, 2000), 119-120, as well as Barry Stroud, "Scepticism, 'Externalism' and the Goal of Epistemology," in *Understanding Human Knowledge* (Oxford: Oxford University Press, 2000), 147-152.

represent its environment. Then it may be that the fact that perception is reliable explains how a creature with a certain set of cognitive faculties manages to represent its environment, without the fact that perception is reliable explaining how human knowledge, in general, is possible. I would like to suggest that Stroud is dissatisfied with externalist explanations, *even if they are correct explanations of the phenomena in question,* because they are not correct explanations of the phenomena *under the description offered.*

Plausibly, the only correct explanation of the phenomena, under the description of interest to the epistemologist, is one given in terms of the irreducible property of *being rationally mandatory.*[14] It would concern the evidence that one has available, and what one is required to believe, given this evidence. Humans, unlike most or all other animals, can recognize that their ordinary beliefs are massively underdetermined by the evidence that they have for them, and can, like Descartes at the beginning of the second meditation, ask what they ought to do about that fact. The under-determination problem is a problem for humans because we are responsive to epistemic reasons in a way that, say, spiders, are not. Asking for an explanation of how *human* knowledge is possible, and asking for it in *those* terms, makes human responsiveness to reasons contextually salient.

An explanation of how human knowledge is possible, which was given in terms of the reliability of human cognitive faculties, would not make essential appeal to the contextually salient property of being responsive to reasons, and would, for that reason, be unsatisfactory. But an explanation of how human knowledge is possible, given in terms of what it is rationally mandatory to believe, does make essential appeal to the fact that humans are responsive to reasons. It would tell us that there are, in fact, propositions that we are rationally required to believe, given the evidence that we have. That an explanation given in terms of the reliability of our cognitive faculties would not make essential appeal to the relevant contextually salient properties is, I believe, the deep reason that Stroud finds externalist accounts of human knowledge unsatisfying.

[14] Although Stroud denies that there is *any* correct explanation. See Stroud, "Understanding Human Knowledge," 120-121.

And if the correct explanation of how human knowledge is possible, when described in those terms, makes essential use of the property of *being rationally mandatory*, then the difficulties that we encountered trying to vindicate the ontological half of the naturalist program will turn out to be significant indeed. For we will have found that there is a property, which cannot be naturalized, but which does essential explanatory work. And that finding would undermine the claims to coherence and comprehensiveness that make naturalism's metaphysical/epistemic picture of the world so attractive.

4. Summary

Popular belief, perhaps, to the contrary, ours is not an anti-metaphysical age. But the most popular metaphysical view is a modest one; it is one that takes its lead from the results of scientific investigation. In fact, this metaphysical view comprises one half of an influential line of thought which also includes a conception of the nature of inquiry in general, and the relation of philosophy to science in particular. I have not argued that this metaphysical view is mistaken, but I have argued that it cannot be vindicated. Any argument for it presupposes that it is false. I went on to suggest that this may be a source of some concern for the broader naturalism of which it is a part. The failure of the reductive program, together with the opacity of explanatory contexts, leaves open the possibility that there are legitimate non-scientific explanations. And, plausibly, how human knowledge is possible, when described in those terms, calls for such an explanation.[15]

Literature

Kornblith, Hilary. *Knowledge and Its Place in Nature*. Oxford: Oxford University Press, 2002.
Mackie, J.L.. *Ethics: Inventing Right and Wrong*. New York: Penguin, 1977.

[15] A draft of this paper was presented as a part of the 2013 Prometheus Seminar Series at Johns Hopkins. I would like to thank my audience for their feedback and encouragement. In addition, thanks are due to Nick Goldberg for providing me with comments on the penultimate draft.

Putnam, Hilary. "Why There Isn't a Ready-Made World." In *Realism and Reason: Philosophical Papers Volume 3*, 205-228. Cambridge: Cambridge University Press, 1983. Originally published in *Synthese* 51 (1982): 141-167.
——— "Beyond the Fact/Value Dichotomy." In *Realism with a Human Face*, edited by James Conant, 135-141. Cambridge: Harvard University Press, 1990. Originally published in *Crítica* 14 (1982): 3-12.
——— *Realism and Reason: Philosophical Papers Volume 3*. Cambridge: Cambridge University Press, 1983.
——— *Realism with a Human Face*. Cambridge: Harvard University Press, 1990.
——— *The Collapse of the Fact/Value Dichotomy, and Other Essays*. Cambridge: Harvard University Press, 2002.
Stroud, Barry. "Understanding Human Knowledge in General." In *Understanding Human Knowledge*, 99-121. Oxford: Oxford University Press, 2000. Originally published in *Knowledge and Skepticism*, edited by Marjorie Klay and Keith Lehrer, 32-50. Boulder, CO: Westview, 1989.
——— "Scepticism, 'Externalism' and the Goal of Epistemology." In *Understanding Human Knowledge*, 139-154. Oxford: Oxford University Press, 2000. Originally published in *Proceedings of the Aristotelian Society: Supplementary Volume* 68 (1994): 291 -307.
——— *The Quest for Reality: Subjectivism & The Metaphysics of Color*. Oxford: Oxford University Press, 2000.
——— *Understanding Human Knowledge*. Oxford: Oxford University Press, 2000.

Quine, Putnam and the Naturalization of Metaphysics

Sander Verhaegh
(University of Groningen)

Naturalists argue that metaphysics ought to be in some sense continuous with science. Putnam has claimed that if we push naturalism to its limits, we have to conclude with Quine that reference is indeterminate. Since Putnam believes Quine's thesis to be extremely counter-intuitive, he regards naturalism to be an unsatisfactory approach to metaphysics. In this essay, I will show that Quine's ideas about reference do not necessarily follow from his naturalism and that, as a result, Putnam's *reductio* argument against naturalism breaks down. In addition, I will argue that Putnam's pluralistic alternative to Quine's views is perfectly compatible with a naturalistic perspective as well and that, in consequence, even for a naturalist the relation between science and metaphysics is less straightforward than it might initially seem to be.

1. Introduction

There is a widespread tendency among present day metaphysicians to think of themselves as 'naturalists'. According to these naturalists, metaphysics ought to be in some sense continuous with science; they endorse the Quinean dictum that the differences between science and metaphysics are "differences only in degree and not in kind".[1] Like W.V. Quine, they stress that we are bound to answer our metaphysical questions within our scientific system of beliefs. We should treat metaphysical questions as scientific questions, accepting "that it is within science itself, and not in some prior philosophy, that reality is to be identified and described".[2]

Prima facie, a thoroughly naturalized metaphysics inevitably seems to lead to an extremely austere physicalistic picture of reality. Indeed, Quine is well known for his taste for "desert landscapes of

[1] Quine, "On Carnap's Views on Ontology," 1951, reprinted in *Ways of Paradox*, rev. ed. (Cambridge, MA: Harvard University Press, 1976), 211.
[2] Quine, "Things and their Place in Theories," in *Theories and Things* (Cambridge, MA: Harvard University Press, 1981), 21.

largely featureless matter" moving to and fro according to the laws of physics.[3] Quine rejects intentional notions as scientifically unacceptable and has developed a naturalistic theory of intentionality in its stead. One of the most radical consequences of this theory is his thesis about *ontological relativity,* which implies that there is no fact of the matter as to what our terms refer to. According to Quine, reference is simply indeterminate by microphysical standards.

Reflecting on Quine's philosophy, Hilary Putnam has playfully remarked that there are two kinds of philosophical genius. On the one hand, there is the genius who proposes "a highly believable [...], coherent, and profound account of reality".[4] On the other hand, philosophical genius can also consist in presenting a system of ideas which is plainly absurd, yet which exhibits fundamental problems for received philosophical views. In the latter case, the importance of the genius in question rests in the deep philosophical puzzles he or she presents for the generations to come. Putnam lists Aristotle and Kant as examples of the first type of genius, while he regards Berkeley as one of the best known examples of the second type; after the work of the latter, philosophers "have been forced" to reconsider their views about both matter and perception. According to Putnam, Quine is a genius of the second type: "just as Berkeley in a sense showed the limits of classical empiricism by pushing it to its logical conclusions, so Quine has shown the limits of naturalism".[5] For Putnam, Quine's ontological relativity is so counter-intuitive that it counts as a strong argument against the naturalism that underlies it: "any doctrine that leads to the conclusion that there is no fact of the matter as to what our terms refer to must be wrong".[6]

In this essay, I will argue that Putnam is mistaken when he rejects naturalism on the basis of this *reductio ad absurdum*. I will show that, although Quine's theory of intentionality is naturalistic in spirit, it does not follow from his naturalism as such. In addition, I will show that Putnam's alternative to Quine's theory is compatible

[3]B. Stroud, "Review Essay: Pursuit of Truth," review of *Pursuit of Truth* by Quine, *Philosophy and Phenomenological Research,* 52:4 (1992): 981.
[4]Putnam, "Quine," *Common Knowledge*, 8:2 (2002): 274.
[5] Ibid., 279.
[6]Putnam, "Reply to Burton Dreben," *Philosophical Topics,* 20:1 (1992): 396.

with a naturalistic approach as well. In doing this, I hope to draw a more realistic picture of what it means to naturalize metaphysics and show that, even for a naturalist, the relation between science and metaphysics is not as straightforward as is often suggested.

2. The Problem of Intentionality

Intentionality is the property of being directed towards some object, whether concrete or abstract, whether existent or not. Our beliefs, our sentences, and our desires, for example, are intentional in that they are always *about* something. According to Franz Brentano, who reinvented the medieval notion at the end of the 19th century, intentionality is a characteristic property of mental phenomena: "[n]o physical phenomenon exhibits anything similar".[7] The problem of intentionality, which has had an enormous impact on 20th century philosophy, is therefore the question of how to reconcile the intentional character of the mental with the physicalistic worldview propagated by the natural sciences.

Probably the most common response to the problem of intentionality is to try to reduce our intentional talk to something non-intentional.[8] Both Quine and Putnam dismiss this option. According to Quine, "there is no breaking out of the intentional vocabulary by explaining its members in other terms".[9] Similarly, Putnam argues that reducing intentionality is impossible because there is "no scientifically describable property that all cases of any particular intentional phenomenon have in common".[10] Yet, although Quine and Putnam agree on the irreducibility of intentional notions, the conclusions that they have drawn from this are diametrically opposed. Quine's solution to the problem of intentionality is to dismiss

[7]Brentano, *Psychology from an Empirical Standpoint,* 1874, translated under the editorship of T. Honderich (London: Routledge and Kegan Paul, 1973), 97.
[8]Some influential examples are F. Dretske, *Knowledge and the Flow of Information* (Cambridge, MA: M.I.T. Press, 1981); R. Millikan, *Language, Thought, and Other Biological Objects* (Cambridge, MA: M.I.T. Press, 1984); and J. Fodor, *Psychosemantics* (Cambridge, MA: M.I.T. Press, 1987).
[9]Quine, *Word and Object* (Cambridge, MA: M.I.T. Press, 1961), 220.
[10]Putnam, *Representation and Reality* (Cambridge, MA: M.I.T. Press, 1988), 2.

intentionality altogether, whereas Putnam's solution is to take intentional talk at face value. In what follows, I will discuss their respective views in turn and argue that naturalism is not the ultimate source of their differences.

3. A Naturalistic Theory of Intentionality

For Quine, the problem of intentionality is a semantic problem; it is the question of how our sentences can be meaningful, the question of how our utterances can be about something. As a naturalist, Quine proposes approaching such semantic matters "in the empirical spirit of natural science".[11] According to Quine, the natural sciences teach us that, in learning our languages, all we have to guide us are the triggerings of our sensory receptors, i.e. the light rays that strike our retinas and the molecules that bombard our eardrums and fingertips.[12] In developing his naturalistic theory of intentionality, Quine seeks to examine the relation between these sensory stimulations and our linguistic utterances.

The major problem with taking sensory stimulations as the basis of a theory of intentionality, however, is that they differ from person to person, whereas our linguistic utterances are intersubjectively understood. Quine solves this problem by proposing that children learn basic observation sentences by being *conditioned* to associate them with the correct set of sensory stimulations, i.e. by being conditioned to use them only on occasions in which this is appropriate according to the other speakers of the language. A child's mastery of the term 'red', for example, can be equated with the "acquisition of the habit of assenting when the term is queried in the presence of red, and only in the presence of red".[13]

[11]Quine, "Philosophical Progress in Language and Language Theory," *Metaphilosophy* 1 (1970): 2.
[12]See Quine, "The Scope and Language of Science," 1954, reprinted in *Ways of Paradox*, 228.
[13]Quine, "The Nature of Natural Knowledge," 1975, reprinted in *Confessions of a Confirmed Extensionalist and Other Essays*, ed. D. Føllesdal and D. Quine (Cambridge, MA: Harvard University Press, 2008), 263.

Now, for our present purposes, it is important to note that if we equate grasping the meaning of a sentence with being able to use the sentence correctly, and Quine does so,[14] it follows that "[t]here is nothing in linguistic meaning beyond what is to be gleaned from overt behavior in observable circumstances".[15] In other words, in Quine's naturalistic theory of intentionality, there is nothing to meaning outside the behavioral facts. The consequence of this position is that Quine rejects the intentional character of meaning altogether:

> One may accept the Brentano thesis either as showing the indispensability of intentional idioms and the importance of an autonomous science of intention, or as showing the baselessness of intentional idioms and the emptiness of a science of intention. My attitude, unlike Brentano's, is the second.[16]

In Quine's theory, the 'aboutness' of sentences and utterances is replaced by a picture in which the meaning of an observation sentence is nothing more than the correlation between certain behavioral facts and certain patterns of sensory stimulation. My utterance of the one-term observation sentence 'Rabbit', for example, is not really 'about' a certain type of object, it is merely an instance of my disposition to utter a certain string of sounds in situations that are similar with respect to the stimulation of my sensory receptors.

4. Ontological Relativity

Quine's naturalistic theory of intentionality has radical consequences for our everyday ideas about reference. One of the consequences that has received much attention in the philosophical literature is Quine's idea that reference is indeterminate. To illustrate this idea, Quine makes use of a thought experiment about a group of linguists who, independently from each other, try to translate the language of some completely unknown tribe into English. According to Quine, it is possible that the linguists come up with partially

[14]See Quine, "Use and Its Place in Meaning," in *Theories and Things*.
[15] Quine, *Pursuit of Truth* (Cambridge MA: Harvard University Press, 1990), 38.
[16]Quine, *Word and Object*, 221.

different conclusions. One could imagine, for example, that one of the linguists concludes that the word 'gavagai', a word that members of the tribe utter whenever a rabbit is present, refers to rabbits, whereas another linguist concludes that 'gavagai' refers to brief temporal segments of rabbit. Both conclusions would arguably be compatible with all publicly observable behavior of members of the tribe. Now, since Quine's naturalistic theory of intentionality implies that publicly observable behavior is all there is to meaning, the conclusion has to be that there is no fact of the matter as to which translation is correct:

> It is meaningless to ask whether, in general, our terms 'rabbit', 'rabbit part', 'number', etc. really refer respectively to rabbits, rabbit parts, numbers, etc., rather than to some ingeniously permuted denotations. It is meaningless to ask this absolutely; we can meaningfully ask it only relative to some background language. When we ask, "Does 'rabbit' really refer to rabbits?" someone can counter with the question: "Refer to rabbits in which sense of 'rabbits'?", thus launching a regress; and we need the background language to regress into. The background language gives the query sense, if only relative sense.[17]

It is important to realize that Quine's thesis about ontological relativity is not epistemological but ontological in character.[18] That is, the point of Quine's thesis is not that it is impossible to determine whether members of the tribe 'actually' refer to rabbits instead of to brief temporal segments of rabbits. Rather, the point is that there is *no fact of the matter* here, there is nothing to discover

[17]Quine, "Ontological Relativity," in *Ontological Relativity and Other Essays* (Cambridge, MA: Harvard University Press, 1969), 48-49. In his later work, Quine often illustrates ontological relativity using *proxy functions*. According to Quine, "[a]ny two ontologies are equally supported by all possible data if we can express a one-to-one correlation, what I call a proxy function, between them". See Quine, "The Growth of Mind and Language," in *Confessions of a Confirmed Extensionalist*, 189. An example of such a proxy function is the function which maps all objects of a theory onto their spatio-temporal complements and which reinterprets all predicates accordingly. Because both the objects and the predicates of the theory are reinterpreted, a sentence like "the rabbit is sitting on the grass" is guaranteed to have the same truth-value as "the complement-rabbit is complement-sitting on the complement-grass".

[18]See M. Friedman, "Physicalism and the Indeterminacy of Translation," *Nous* 9:4; and R. Gibson, "Translation, Physics, and Facts of the Matter," in: *The Philosophy of W.V. Quine*, ed. L. Hahn and P. Schilpp (La Salle: Open Court, 1986).

because there is nothing to reference beyond the behavioral facts.[19] What seems quite certain is that 'gavagai' probably does not refer to dogs, for it seems very difficult to come up with such a translation that does justice to the behavioral facts.[20] What is clear, however, is that all publicly observable behavior, which is all there is to meaning, according to Quine, is not enough to the make reference completely determinate.

5. Putnam's *Reductio*

Putnam believes that Quine is right in claiming that a naturalistic approach to intentionality leads to the conclusion that reference is indeterminate. In fact, Putnam has constructed a model-theoretic argument that leads to more or less the same conclusion.[21] Still, Putnam and Quine draw diametrically opposed conclusions. Where Quine accepts ontological relativity as a consequence of his naturalistic theory of intentionality, Putnam holds that "any doctrine that leads to the conclusion that there is no fact of the matter as to what our terms refer to must be wrong".[22] And since Putnam believes that there is "not [...] any hope" that one can coherently accept naturalism "without being driven to [...] ontological relativity",[23] he regards naturalism to

[19]See Quine, *Word and Object*, 73.

[20]Igor Douven shows that there are indeed limits to ontological relativity. According to Douven, one cannot reinterpret a theory's ontology as drastically as one pleases without violating the behavioral facts. Whereas the truth-conditions of a sentence are preserved in reinterpreting one's ontology in terms of another, the action conditions of a sentence and the plausibility ordering among sentences might also influence the behavior of a subject, and therefore might constrain the choice of the ontology for a theory. See Douven, "Empiricist Semantics and Indeterminism of Reference," in *Quine: Naturalized Epistemology, Perceptual Knowledge and Ontology*, ed. L. Decock and L. Horsten (Amsterdam: Atlanta, 2000).

[21]See Putnam, *Reason, Truth, and History* (Cambridge, MA: Cambridge University Press, 1981), Ch. 2; and Putnam, "Model Theory and the "Factuality" of Semantics," 1989, reprinted in *Words and Life*, ed. J. Conant (Cambridge, MA: Harvard University Press, 1994).

[22]Putnam, "Reply to Burton Dreben," 396.

[23]Putnam, "Quine," 279.

be an absurd position: "another philosopher's *modus ponens* is my *modus tollens*".[24]

In the later stages of his career, Putnam has come to defend an alternative to Quine's views, which he himself has called (conceptual) *pluralism*. This pluralism can be characterized as

> the insistence that the various sorts of statements that are regarded as less than fully rational discourse, as somehow of merely 'heuristic' significance, by one or another of the 'naturalists' [...] are *bona fide* statements, "as fully governed by norms of truth and validity as any other statements".[25]

Conceptual pluralism, in other words, is the claim that there is not one vocabulary (e.g. the vocabulary of physics) that has an ontologically or epistemologically exclusive status. According to Putnam, the problem with Quine's views is that they are still guided by outdated notions of 'objectivity', 'existence', and 'factuality'; notions that also cause Quine to give up on intentionality.

6. Regimentation

Before I turn to Putnam's pluralism, however, let me briefly discuss these so-called outdated notions of 'objectivity', 'existence', and 'factuality'. In Quine's naturalized metaphysics, ontology is the study of which entities we are committed to on the basis of our best scientific theories. The naturalized metaphysician tries to determine the minimal set of entities we need to posit to do justice to our scientific beliefs. The problem with this project, however, as Putnam notes, is that "science does not wear its philosophical significance on its sleeve".[26] Whether or not science commits us to elementary particles, tables, numbers, and/or psychological states such as beliefs, for example, is not something that can be precisely determined by reading numerous scientific papers. The naturalized philosopher

[24]Putnam, "Meaning Holism," in *The Philosophy of W.V. Quine*, 425.
[25]Putnam, "The Content and Appeal of "Naturalism"," in *Naturalism in Question*, ed. M. de Caro and D. Macarthur (Cambridge, MA: Harvard University Press), 61. The sub-sentence between quotation marks is a reference to Conant, "Wittgenstein's Philosophy of Mathematics," *Proceedings of the Aristotelian Society* 97:2 (1997): 202.
[26]Putnam, "Quine," 273.

therefore needs a criterion that determines which minimal set of entities is to be posited to reflect our scientific knowledge in the best possible way.

For Quine, the solution to this problem is to reformulate our scientific theories into the language of classical first-order logic with identity, a process that he calls *regimentation*. The advantage of regimenting scientific theories in this way is that in first-order logic both ontological commitments and inferential relations are precisely defined. According to Quine, "entities of a given sort are assumed by a [regimented] theory if and only if some of them must be counted among the values of the variables in order that the statements affirmed in the theory be true".[27] This is Quine's criterion of ontological commitment, and it is often summarized in the slogan "to be is to be a value of a bound variable".

Quine's naturalized ontology, therefore, implies that only entities that are the values of the bound variables in a regimented version of our scientific theories are to be classified as real. And according to Quine, we only need to posit two types of entities in this way, viz. physical objects and mathematical sets.[28] All other entities, whatever their usefulness "in the market place or in the laboratory",[29] are dispensable for theoretical purposes; Quine cannot accord them the same ontological status as the entities included in his regimented theory of the world. This is why Quine often speaks about these other entities as "half-entities in a second-grade system"[30] or as "Grade B idiom".[31] Now, consider the following three statements:

(1) Adam refers to the rabbit in front of him whenever he uses the term 'gavagai'.
(2) Adam refers to the temporal segments of rabbit in front of him whenever he uses the term 'gavagai'.

---

[27]Quine, "Logic and the Reification of Universals," in *From a Logical Point of View*, 2nd ed. (London: Oxford University Press, 1961), 103.
[28]See Quine, "The Scope and Language of Science," 244.
[29]Quine, *Word and Object*, 228.
[30]Quine, "Speaking of Objects," in *Ontological Relativity,* 23-5.
[31]Quine, "Reply to Davidson," in *Words and Objections*, ed. D. Davidson and J. Hintikka (Dordrecht: Reidel, 1969), 335.

(3) Adam has the disposition to utter or assent to 'gavagai' whenever there are rabbits in his visual field.

According to Quine, one can only say of statements like (3) that they are 'true', 'factual' and 'objective', because only statements like (3)—once reformulated in the language of classical first-order logic—would be part of his regimented theory of the world. The difference between the statements (1) and (2) simply transcends the behavioral and hence the physical facts:

> I speak as a *physicalist* in saying there is no fact of the matter [...] behavioral criteria that will ordinarily decide in favor of one translation rather than another. When they do, there is emphatically a fact of the matter *by microphysical standards*; for clearly any difference in overt behavior, vocal or otherwise, reflects extravagant differences in the distribution of elementary physical states.[32]

The problem with (1) and (2), in other words, is that the difference between them does not seem to be reducible to differences in elementary physical states. And since his physicalism commits Quine to the thesis that "nothing happens in the world, not the flicker of a thought, without some redistribution of microphysical states",[33] he has to conclude that one cannot apply norms like truth, factuality or objectivity to statements like (1) and (2).

7. Putnam's Alternative: Pluralism

According to Putnam, on the other hand, none of the statements (1)-(3) can be deprived of qualifications like 'factual', 'objective' or 'true', just because they are stated with one vocabulary or another. According to Putnam, "ontological approaches to 'demarcation' are bankrupt"; it is absurd to suppose that we should dismiss so many of

[32] Quine, "Facts of the matter," in *Confessions of a Confirmed Extensionalist,* 284, my italics.
[33] Quine, "Goodman's Ways of Worldmaking," *Theories and Things*, 98.

our beliefs about what sort of statements can be true on the basis of some misguided ontological criteria.[34]

Still, Putnam agrees with Quine that there can be no change without physical change. According to Putnam, however, this does not imply that all facts are physical facts.[35] Ironically, Putnam's justification of this position depends on one of Quine's theses, viz. the idea that there is no strict distinction between the factual and the conventional aspects of a statement.

In his "Two Dogmas of Empiricism", Quine rejects the idea that we can test our scientific statements about the world individually. Rather, "our statements about the external world face the tribunal of sense experience [...] only as a corporate body".[36] The implication of this view is that one cannot logically determine which statement in this corporate body is falsified whenever an observation conflicts with our predictions. Quine famously concluded from this that "[a]ny statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system" and that "no statement is immune to revision".[37] The implication here is that statements that are thought to be completely factual might also be 'held true come what may' and that statements that are thought to be completely conventional might also be revised in the light of adverse experience.

Putnam, who has always agreed with Quine on this point,[38] has made important contributions towards providing historical evidence for Quine's thesis by pointing at some historical cases in which conventionally introduced definitions were, in fact, revised in the light

[34]See Putnam, "The Idea of Science," in *Words and Life*, 489; and Putnam, *The Threefold Cord: Mind, Body, and World* (New York: Columbia University Press, 1999), 7.

[35]Putnam, "Meaning Holism," 424-5.

[36]Quine, "Two Dogmas of Empiricism," 1951, reprinted in *From a Logical Point of View*, 41.

[37]Ibid., 43.

[38]"[W]hile there is an aspect of conventionality and an aspect of fact in everything we say that is true, we fall into hopeless philosophical error if we commit a "fallacy of division" and conclude that there must be a part of the truth that is the "conventional part" and a part that is the "factual part"." See Putnam, preface to *Realism with a Human Face*, ed. J. Conant (Cambridge, MA: Harvard University Press, 1990), x.

of conflicting empirical evidence.[39] Now, according to Putnam, granting that statement (3) might be true, objective and factual, while denying that statements like (1) and (2) can be classified in the same way is imposing a rigid distinction, whereas in fact such a distinction cannot be made. Every statement has a conventional component and every statement has a factual component, even if the factual component in a statement (3) might be larger.

Putnam illustrates this graduality by arguing that most people would probably agree on the following ranking of properties from being more subjective to being more objective:[40]

(a) Being very amusing.
(b) Being a match that *would have* lit *if* it *had been* struck.
(c) Being an utterance that means 'Do you speak French?' in English.
(d) Being soluble.
(e) Being a region of space which contains at least one hydrogen atom (in classical physics).

For Putnam the problem with Quine's ontological views is that they impose a strict dichotomy between statements that can be true, objective, and factual and statements that cannot:

> Quine [...] would put the cut between [(c) and (d)]—counting both dispositional predicates (such as 'soluble') and non-dispositional predicates from fundamental physics as 'objective' and all the others as more or less subjective [...] my own view [...] is that the enterprise isn't worth the candle. The game is played out. We can make a rough sort of rank ordering (although even here there are disagreements), but the idea of a 'point at

[39]See, for example, Putnam, "The Analytic and the Synthetic," 1962, reprinted in *Mind, Language, and Reality: Philosophical Papers,* Volume 2 (Cambridge: Cambridge University Press, 1975); and Putnam, "The Refutation of Conventionalism," in *Mind, Language, and Reality*. An influential example is Putnam's description of the history of the concept of "momentum" in physics. Putnam shows that although "momentum" was conventionally introduced into physics as "mass times velocity," Einstein established that this equation had to be revised: the momentum of a particle only approaches mass times velocity at non-relativistic speeds. As a result "the statement that momentum is *exactly* equal to mass times velocity was revised [...] *the statement that was originally a 'definition"*. See Putnam, *Representation and Reality*, 9-10.
[40]Putnam, *The Many Faces of Realism* (La Salle: Open Court, 1987), 40.

> which' subjectivity ceases and Objectivity-with-a-capital-O begins has proved chimerical.[41]

Putnam's objection to Quine can thus be stated as the complaint that although Quine has shown that the fact-convention dichotomy cannot be maintained, he does not recognize that giving up this distinction also has implications for the way in which he can use notions like 'existence', 'objectivity', and 'factuality'. For Putnam "the idea that there is an Archimedean point, or a use of 'exist' inherent in the world itself [...] is an illusion".[42]

8. Varieties of Naturalism

Given that Putnam's main argument for pluralism is based on Quine's rejection of the fact-convention distinction, one may wonder why Quine has not drawn the same conclusions as Putnam. It is my contention that Quine could have drawn the same conclusions, but that he did not because of pragmatic considerations. To show this, we will have to dig deeper into Quine's views on the nature of regimentation.

Earlier, I described regimentation as the process of reformulating or rephrasing our scientific theories into the language of classical first-order logic. What I did not discuss, however, is what Quine means by 'reformulating' and 'rephrasing'. Because Quine rejects notions like 'sameness of meaning', he cannot identify these terms with creating a sentence-by-sentence translation synonymously with ordinary language. Rather, for Quine, regimentation is a creative enterprise:

> [R]egimentation is not a matter of eliciting some latent but determinate ontological content of ordinary language. It is a matter rather of *freely creating* an ontology-oriented language that can supplant ordinary language in serving some particular purposes that one has in mind.[43]

[41]Ibid., 28.
[42]Ibid., 20.
[43]Quine, "Facts of the matter," 285, my italics.

According to Quine, the naturalized metaphysician is free in choosing whether to regiment our scientific beliefs in one way or another. The *explicandum* and the *explicans* must roughly fulfill the same function in the overall theoretical structure, but "[b]eyond those conditions of partial agreement, dictated by our interests and purposes, any traits of the explicans come under the head of 'don't-cares'".[44]

Quine's open-mindedness with regard to regimentation is illustrated in the way in which he justifies the choices he makes in the regimentation process. With respect to the question whether his regimented system should contain a two-valued or a many-valued truth-predicate, for example, he argues that although opting for bivalence comes "at no small price" due to phenomena like vagueness, he still chooses a two-valued logic "for the *simplicity* of theory that it affords".[45] Similarly, in pondering whether to adopt a different logic in order to account better for the reasoning involved in quantum mechanics, Quine claims that we should not "underestimate the price of a deviant logic. There is a serious loss of *simplicity* [...] The price is not quite prohibitive, but the returns had better be good".[46]

Quine thus seems mostly guided by pragmatic values in regimenting our scientific knowledge: "the general principle is that we should choose our canonical notation in a way which maximizes the simplicity, clarity, and efficacy of our system of knowledge as a whole. For Quine this is all that matters—there are no other constraints."[47] The problem with pragmatic values, however, is (1) that there is no generally agreed upon list of pragmatic values which regimented scientific theories should adhere to,[48] and (2) that it is completely unclear how we should balance all these different virtues. This implies that different naturalists will not only make different choices in regimenting our scientific theories of the world, they will

[44]Quine, *Word and Object*, 259.
[45]Quine, "What Price Bivalence?" in *Theories and Things*, 32, my italics.
[46]Quine, *Philosophy of Logic* (Englewood Cliffs: Prentice-Hall, 1970), 86.
[47]Hylton, *Quine* (New York: Routledge, 2007), 259-60.
[48]Quine himself gives different lists of pragmatic virtues in different publications. Cf. Quine, "Posits and Reality," 1955, reprinted in *Ways of Paradox*, 247; Quine, *From Stimulus to Science* (Cambridge, MA: Harvard University Press, 1995), 49; and Quine and Ullian, *The Web of Beliefs,* 2nd ed. (New York, Random House, 1978), Ch. 6.

also have different desiderata that their regimentations should fulfill; how one regiments a certain notion depends on "the particular purposes that one has in mind".[49] One naturalist will use one set of pragmatic values with a particular system of weighing those values, whereas another naturalist can emphasize a relatively different set of pragmatic values and a relatively different weighing system.

9. Pluralistic Naturalism

Quine himself is clearly guided by pragmatic principles like simplicity, clarity, and systematicity; hence his choice to limit his regimented theory to physics and the necessary mathematical structures that come with it. One could imagine a different naturalist, however, let us call him a *pluralistic naturalist*, who would place much more emphasis on pragmatic values such as comprehensiveness, familiarity, and maintaining the richness of ordinary language. Would Quine regard the choices of such a pluralistic naturalist as somehow inferior? Probably he would, but the point of the above argument is that he would not reject this alternative on the basis of naturalistic criteria. Rather, his dismissal would be guided by his "taste for desert landscapes",[50] by his pragmatic conviction that naturalized metaphysicians ought to construct a framework that connects our scientific theories as simply and clearly as possible.

So, to come back to the question with which we started the previous section: why does Quine not draw the same conclusions as Putnam, given that they both reject the fact-convention dichotomy? That is, how can Quine impose a strict distinction between statements that are true, objective and factual, and statements that are not? In my opinion, the above discussion shows that Quine only draws this distinction on pragmatic grounds. Quine does not think that he is in the business of describing the world as it is in itself, for, as a naturalist, he does not think that the question as to which entities *really* exist makes sense; there is no extra-scientific notion of existence

[49]Quine, "Facts of the matter," 285.
[50]Quine, "On What There Is," 1948, reprinted in *From a Logical Point of View*, 4.

that we can discuss from "a vantage point outside our conceptual scheme",[51] just as Putnam claims that "the idea that there is an Archimedean point, or a use of 'exist' inherent in the world itself [...] is an illusion".[52] Quine just *chooses* to adopt a physicalistic notion of 'existence', 'objectivity', and 'factuality' in his regimented theory because this results in a simpler and more systematic theory of the world.

As a result, Putnam's *reductio* argument breaks down. Naturalism does not imply that we should distinguish between a first and a second-grade system of entities, nor does it imply that reference is indeterminate. Only, if we regiment our scientific theories in such a way as to obtain maximal simplicity and clarity (although Putnam would probably disagree that the result is really a simple and clear theory of the world), naturalism leads to physicalism and hence to Quine's views on intentionality and ontology. Yet, simplicity, systematicity and clarity are not the only possible desiderata in our regimentation practices, nor do we have to put so much weight on these pragmatic virtues. This means that Putnam's pluralistic view is compatible with a naturalistic approach to philosophy as well. We could imagine a pluralistic naturalist who places much more emphasis on pragmatic virtues such as comprehensiveness, familiarity, and maintaining the richness of ordinary language.

Is this 'pluralistic naturalism' really a variant of naturalism? I believe it is. Putnam's metaphysics also greatly respects scientific practice. Putnam only rejects the idea that we can decide "what is and what is not 'science' [...] on the basis of [bankrupt] ontological criteria".[53] According to Putnam, there is no such thing as "Science-with-a-capital-S", or "The Scientific Method". Rather, what we commonly call science consists of a great variety of different scientific practices. "Why", Putnam asks us, should we "expect the sciences to have more than a family resemblance to one another? [...] there is no set of 'essential' properties that all the sciences have in common".[54]

---

[51]Quine, *Word and Object*, 275.
[52]Putnam, *The Many Faces of Realism*, 20.
[53]Putnam, "The Idea of Science," 1990, reprinted in *Words and Life*, 489.
[54]Putnam, "The Diversity of the Sciences," 1987; reprinted in *Words and Life*, 471-2.

Putnam's pluralism, then, is perfectly compatible with a naturalistic approach to philosophy as long as 'science' is also interpreted pluralistically. Where Quine's naturalism can be described as *monistic,* in that he decides to take only the language of physics at face value, Putnam's position can be described as a *pluralistic naturalism*, as the view that for any word *L*, even if *L* is the intentional vocabulary, it holds that questions with respect to *L* are not answerable to any tribunal outside *L* and not in need of any justification beyond the methods used within *L*.[55]

As a result, it is a mistake to believe that Putnam's pluralism is based on a dismissal of naturalism. For Putnam's position itself can be described as naturalistic as well. In fact, even Putnam's objection to Quine can be regarded as naturalistic in spirit. For, according to Putnam, anyone who denies the pluralistic character of science and who argues that all science should be like physics, confuses "respect for science with uncritical acceptance of a materialist ideology".[56] A worldview that is thoroughly physicalistic, i.e. a view in which intentionality is rejected because there is no fact of the matter to questions about meaning and reference by physical standards, simply leaves too much out of the picture. It does not do credit to the richness and plurality of the descriptions and explanations that science actually provides. Although Quine claims to have a broad conception of science,[57] the starting point of his theory of intentionality is that we have to "approach semantical matters in the empirical spirit of *natural* science".[58] Why not approach questions about intentionality from above, i.e. "from the standpoint of [our]

[55] Intermediate positions between *monism* and *pluralism* are also possible. Penelope Maddy's 'mathematical naturalism', for example, could be described as *dualistic* in that she only makes a strict distinction between scientific and mathematical vocabulary: "where Quine holds that science is 'not answerable to any supra-scientific tribunal, and not in need of any justification beyond observation and the hypothetico-deductive method' [...] the mathematical naturalist adds that *mathematics* is not answerable to any extra-mathematical tribunal and not in need of any justification beyond proof and the axiomatic method". See Maddy, *Naturalism in Mathematics* (Oxford: Oxford University Press, 1997), 184.

[56] Putnam, Preface to *Word and Life*, vi.

[57] See Quine, *From Stimulus to Science,* 49.

[58] Quine, "Philosophical Progress in Language and Language Theory," 8, my italics. Cf. S. Haack, *Evidence and Inquiry* (Oxford: Blackwell Publishers Ltd., 1993), Ch. 6.

intentional notions"?[59] After all, interpretation is a key methodological tool in many of the 'softer' sciences, that is, sciences that will not improve in the future by imitating physics.[60] Rather than claiming that Quine has shown us the limits of naturalism, just as Berkeley has shown us the limits of classical empiricism, Putnam could have claimed that Quine's views are not naturalistic enough.

10. Conclusion

I conclude that, in the end, the differences between Quine and Putnam largely depend on their different pictures of science. According to Quine, the different sciences should be respected, but physics is the ultimate arbiter of truth, factuality and existence. Whenever some practices cannot be reduced to physics—as is the case with our intentional vocabulary—we have to reject them and treat them as second-grade. What remains are the behavioral facts. This ultimately leads to ontological relativity. For Putnam, such an account leaves too much out and does not do justice to the highly holistic and normative character of our interpretative practices. His solution to the non-reducibility of the intentional is to take our intentional talk at face value. Certainly there is no change without physical change, but this does not mean that all facts are physical facts. As a result, there is not one 'Science' that sets the standard for the rest. Rather, Putnam argues, science is a plurality of strongly interconnected, yet ultimately independent practices that all should be judged on the basis of criteria internal to those practices themselves.

In this essay, I have argued that both pictures are naturalistic in spirit and can be defended on pragmatic grounds. An advantage of Putnam's pluralism is that it does justice to the richness and diversity of the sciences. Quine's variant of naturalism, on the other hand, does more justice to the drive for unity, precision and simplicity that is also characteristic of the scientific enterprise. I will leave it to the reader to decide which of the two pictures is more plausible or does better

[59]Putnam, "Realism without Absolutes," 1993, reprinted in *Words and Life*, 290.
[60]Putnam, "The Idea of Science," 490.

justice to our scientific practices.[61] As the choice between the two pictures is largely pragmatic, one's choice will probably not only depend on one's view of what science actually is but also on one's ideas about what science ought to be.[62]

Literature

Brentano, F. *Psychology from an Empirical Standpoint*, 1874, translated under the editorship of T. Honderich. London: Routledge and Kegan Paul, 1973.

Conant, J. "Wittgenstein's Philosophy of Mathematics." *Proceedings of the Aristotelian Society* 97:2 (1997): 195-222.

Douven, I. "Empiricist Semantics and Indeterminism of Reference." In *Quine: Naturalized Epistemology, Perceptual Knowledge and Ontology*, edited by L. Decock and L. Horsten, 171-88. Amsterdam: Atlanta, 2000.

Dretske, F. *Knowledge and the Flow of Information*. Cambridge, MA: M.I.T. Press, 1981.

Fodor, J. *Psychosemantics*. Cambridge, MA: M.I.T. Press, 1987.

Friedman, M. "Physicalism and the Indeterminacy of Translation." *Nous* 9:4 (1975): 353-74.

Gibson jr., R. "Translation, Physics, and Facts of the Matter." In Hahn and Schilpp, *The Philosophy of W.V. Quine*, 139-54.

Haack, S. *Evidence and Inquiry*. Oxford: Blackwell Publishers Ltd., 1993.

Hahn, L. and P. Schilpp. *The Philosophy of W.V. Quin*e. La Salle: Open Court, 1986.

Hylton, P. *Quine*. New York: Routledge, 2007.

Maddy, P. *Naturalism in Mathematics*. Oxford: Oxford University Press, 1997.

Millikan, R. *Language, Thought, and Other Biological Objects*. Cambridge, MA: M.I.T.Press, 1984.

Putnam, H. "The Analytic and the Synthetic," 1962. In *Mind, Language, and Reality*, 33-69.

———. *Mind, Language, and Reality: Philosophical Papers*, volume 2. Cambridge: Cambridge University Press, 1975.

[61]Of course, the reader also has a third option: He or she might decide to reject some of the ideas on which Quine and Putnam agree and that have been presupposed in this essay, e.g. their claim that intentional notions cannot be reduced to non-intentional notions, their claim that there is no strict fact-convention dichotomy, or their claim that first-philosophy ought to be rejected.

[62]I thank Igor Douven, Jeanne Peijnenburg, Jan-Willem Romeijn, Allard Tamminga, the members of the Groningen research colloquium in theoretical philosophy, the members of the Work in Progress colloquium, and the people who attended my talks at the Universities of Milan and Groningen for their comments. This research is funded by the Netherlands Organization for Scientific Research (NWO), Grant 322-20-001.

———. "The Refutation of Conventionalism," 1975. In *Mind, Language, and Reality*, 152-91.

———. *Reason, Truth, and History*. Cambridge, MA: Cambridge University Press, 1981.

———. "Meaning Holism." 1986. In Hahn and Schilpp, *The Philosophy of W.V. Quine*, 405–26.

———. *The Many Faces of Realism*. La Salle: Open Court, 1987.

———. "The Diversity of the Sciences," 1987. In *Words and Life*, 463-80.

———. *Representation and Reality*. Cambridge, MA: M.I.T.Press, 1988.

———. "Model Theory and the "Factuality" of Semantics," 1989. In *Words and Life*, 351-75.

———. *Realism with a Human Face*, edited by J. Conant. Cambridge, MA: Harvard University Press, 1990.

———. "The Idea of Science," 1990. In *Words and* Life, 481-91.

———. "Reply to Burton Dreben." *Philosophical Topics* 20:1 (1992): 393–99.

———. "Realism without Absolutes," 1993. In *Words and Life*,279-94.

———. *Words and Life*, edited by J. Conant. Cambridge, MA: Harvard University Press, 1994.

———. *The Threefold Cord: Mind, Body, and World*. New York: Columbia University Press, 1999.

———. "Quine." *Common Knowledge* 8:2 (2002): 273–9.

———. "The Content and Appeal of 'Naturalism'." In *Naturalism in Question*, edited by M. de Caro and D. Macarthur, 59-70. Cambridge, MA: Harvard University Press, 2004.

Quine, W.V. "On What There Is," 1948. In *From a Logical Point of View*, 1-19.

———. "Two Dogmas of Empiricism," 1951. In *From a Logical Point of View,* 20-45.

———. "On Carnap's Views on Ontology," 1951. In *Ways of Paradox*, 203-11.

———. "The Scope and Language of Science," 1954. In *Ways of Paradox,* 228-45.

———. "Posits and Reality," 1955. In *Ways of Paradox,* 246-54.

———. "Speaking of Objects," 1957. In *Ontological Relativity*, 1-25.

———. *Word and Object*. Cambridge, MA: M.I.T. Press, 1961.

———. *From a Logical Point of View,* 2nd ed. London: Oxford University Press, 1961.

———. "Logic and the Reification of Universals," 1961. In *From a Logical Point of View,* 102-29.

———. *Ontological Relativity and other essays*. Cambridge, MA: Harvard University Press, 1969.

———. "Ontological Relativity," 1969. In *Ontological Relativity*, 26-68.

———. "Reply to Davidson." In *Words and Objections*, edited by D. Davidson and J. Hintikka, 333-5. Dordrecht: Reidel, 1969.

———. *Philosophy of Logic*. Englewood Cliffs: Prentice-Hall, 1970.

———. "Philosophical Progress in Language and Language Theory." *Metaphilosophy* 1 (1970): 2–19.

———. "The Nature of Natural Knowledge," 1975. In *Confessions of a Confirmed Extensionalist*, 257-70.

———. *The Ways of Paradox,* rev. ed. Cambridge, MA:Harvard University Press, 1976.

———. "Facts of the matter," 1977. In *Confessions of a Confirmed Extensionalist*, 271-86.

———. "Goodman's Ways of Worldmaking," 1978. In *Theories and Things*, 96-99.

———. *Theories and Things*. Cambridge, MA: Harvard University Press, 1981.

———. "Things and Their Place in Theories," 1981.In *Theories and Things*, 1-23.

———. "What Price Bivalence?" 1981. In *Theories and Things,* 31-7.

———. "Use and Its Place in Meaning." In *Theories and Things*, 43-54.

———. *Pursuit of Truth*. Cambridge, MA: Harvard University Press, 1990.

———. *From Stimulus to Science*. Cambridge, MA: Harvard University Press, 1995.

———. "The Growth of Mind and Language," 1997. In *Confessions of a Confirmed Extensionalist*, 182-91.

———. *Confessions of a Confirmed Extensionalist and Other Essays*, edited by D. Føllesdal and D. Quine. Cambridge, MA: Harvard University Press, 2008.

———. and J.S. Ullian. *The Web of Beliefs*, 2nd ed. New York: Random House, 1978.

Stroud, B. "Review Essay: Pursuit of Truth." *Philosophy and Phenomenological Research* 52:4 (1992): 981-7.

University
of Bamberg
Press

This collection of essays originated in the Summer School 'Metaphysics or Modernity?', which was held at the University of Bamberg in August 2012. Designed as a forum for graduate students in philosophy, the Summer School brought together a highly diverse group of young academics who – more often than not – came from utterly different schools and traditions of thought.

This pluralism is reflected in the pages of this book. While the volume is roughly divided into two halves – one with a more historical focus, the other with a more systematic focus – the reader will find an unusually wide array of topics and questions treated here. Since the aforementioned pluralism was one of the main strengths of our Summer School, this is something in which we take much pride.

eISBN 978-3-86309-181-1

www.uni-bamberg.de/ubp