

SORITES

σωρίτης

An International Electronic Magazine of Analytical Philosophy

Indexed and Abstracted in *THE PHILOSOPHER'S INDEX*

ISSN 1135-1349

Legal Deposit Registration: M 14867-1995

Editor: Lorenzo Peña

.....
Associate Editor: Txetxu Ausín

(Spanish Institute for Advanced Studies)

Board of Editorial Consultants:

Jean-Yves Béziau, Enrique Alonso, Guillermo Hurtado, Manuel Liz,
Raymundo Morado

Regular-Mail Address:

Prof. Lorenzo Peña

CSIC [Spanish Institute for Advanced Studies]

Department of Theoretical Philosophy

Pinar 25

E-28006 Madrid

Spain

Fax +3491 564 52 52

Voice Tph +3491 411 70 60, ext 18

INTERNET ACCESS:

<<http://www.sorites.org/>>

<<http://www.ifs.csic.es/sorites/>>

Editorial e-mail inbox: <sorites@sorites.org>

Inquiries and subscription-requests: <sorites@sorites.org>

Issue #16 — December 2005

SORITES (ΣΩΡΙΤΗΣ)

ISSN 1135-1349

Issue #16 — December 2005

Copyright © by SORITES and the authors

MAIN INTERNET ACCESS:

<<http://www.sorites.org/>>

<<http://www.ifs.csic.es/sorites/>>

<sorites@sorites.org> (Editorial e-mail inbox, esp. for submissions)

<sorites@sorites.org> (Inquiries and subscription-requests)

SORITES

ISSN 1135-1349

ROLL OF REFEREES

Rainer Born Johannes-Kepler Universitaet Linz (Austria)
Amedeo Conte University of Pavia (Italy)
Newton C.A. da Costa University of São Paulo (Brazil)
Marcelo Dascal University of Tel Aviv (Israel)
Dorothy Edgington Birbeck College (London, UK)
Graeme Forbes Tulane University (New Orleans, Louisiana, USA)
Manuel García-Carpintero University of Barcelona (Spain)
Laurence Goldstein University of Hong Kong (Hong Kong)
Jorge Gracia State University of New York, Buffalo (USA)
Nicholas Griffin McMaster University (Hamilton, Ontario, Canada)
Rudolf Haller Karl-Franzens-Universitaet Graz (Austria)
Terence Horgan University of Memphis (Tennessee, USA)
Victoria Iturralde Univ. of the Basque Country (San Sebastian, Spain)
Tomis E. Kapitan Northern Illinois University (USA)
Manuel Liz University of La Laguna (Canary Islands, Spain)
Peter Menzies Australian National University (Canberra, Australia)
Carlos Moya University of Valencia (Spain)
Kevin Mulligan University of Geneva (Switzerland)
Jesús Padilla-Gálvez Johannes-Kepler Universitaet Linz (Austria)
Philip Pettit Australian National University (Canberra, Australia)
Graham Priest University of Queensland (Brisbane, Australia)
Eduardo Rabossi University of Buenos Aires (Argentina)
David-Hillel Ruben School of Oriental and African Studies, University of London
Mark Sainsbury King's College (London, UK)
Daniel Schulthess University of Neuchâtel (Switzerland)
Peter Simons University of Leeds (Leeds, UK)
Ernest Sosa Brown University (Providence, Rhode Island, USA)
Friedrich Stadler Institut «Wien Kreis» (Vienna, Austria)

SORITES

ISSN 1135-1349

Issue #16 — December 2005

TABLE OF CONTENTS

— Abstracts of the Papers	03
— «Wittgenstein, Consciousness, and the Mind» by Pär Sundström . . .	06
— «The Mereology of Events» by Robert Allen	23
— «Dismantling the Straw Man: An Analysis of the Arguments of Hume and Berkeley Against Locke's Doctrine of Abstract Ideas» by Rhys McKinnon	38
— «Tarskian Metamathematics in Carnap's Metalogic» by Jesús Padilla-Gálvez	46
— «Why Axiomatize Arithmetic?» by Charles Sayward	54
— «Is Theism More Rational Than Agnosticism: A Critique of Arguments for the Necessary Existence of God?» by David Kimweli	62
— «Rules and Realism: Remarks on the Poverty of Brute Facts» by J. Jeremy Wisnewski	74
— «What is a Value Judgement?» by Georg Spielthener	82
— «Hyper Libertarianism and Moral Luck» by Gerald K. Harrison	93
— Copyright Notice and Legal Disclaimer	103
— Release Notice	105

SORITES (ΣΩΡΙΤΗΣ), ISSN 1135-1349
<http://www.sorites.org>
Issue #16 — December 2005. Pp. 3-5
Abstracts of the Papers
Copyright © by SORITES and the authors

ABSTRACTS OF THE PAPERS

Wittgenstein, Consciousness, and the Mind

by Pär Sundström

Contrary to philosophical tradition, modern theorists of the mind have often downplayed the importance of consciousness. Instead, they have accounted for the mind in terms of phenomena like mechanisms, dispositions, abilities and even environmental features. One of many inspirations for this trend is a series of passages of the later Wittgenstein. These passages discuss a variety of specific mental phenomena, like searching, comparing, understanding and reading. The passages have often been taken them to show that one may exemplify any of the phenomena at issue without being in any particular type of conscious state. I claim that the passages do not support this conclusion, and that the conclusion is, arguably, false. My conclusion is that consciousness may be a more important aspect of the mind than is supposed by many contemporary theorists—both Wittgensteinians and others.

The Mereology of Events

by Robert Allen

I demonstrate here that it is possible for an event to be identical with one of its proper parts, refuting the key premise in Lawrence Lombard's argument for the essentiality of an event's time. I also propose and defend an alternative to his criterion of event identity.

Dismantling the Straw Man: An Analysis of the Arguments of Hume and Berkeley Against Locke's Doctrine of Abstract Ideas

by Rhys McKinnon

Many believe that George Berkeley (*Principles of Human Knowledge*) and, subsequently, David Hume (*A Treatise of Human Nature*) offer devastating arguments against John Locke's (*An Essay Concerning Human Understanding*) theory of abstract ideas. It is the purpose of this paper to clarify the attacks given a close reading of Locke. It will be shown that many of the arguments of Berkeley and Hume are of a straw man nature and, moreover, that some of their conclusions are actually in accord with Locke.

Tarskian Metamathematics in Carnap's Metalogic

by Jesús Padilla Gálvez

This paper examines how the intellectual heritage of Tarski's «scientific semantics» contributes to the metalogic of Carnap, and viceversa. It seeks first to establish the connections between the Warsaw School in metamathematics and semantics. Secondly, the author explores the relationship between the Warsaw School and the Vienna Circle. Third, the specific influences of Tarski's program on the logical syntax of language will be analyzed and finally, the internal discussions conducted within the Vienna Circle in relation to Carnap's contribution to metalogic will be discussed.

Why Axiomatize Arithmetic?

by Charles Sayward

This is a dialogue in the philosophy of mathematics. The following issues are discussed: Are the Peano axioms for arithmetic epistemologically irrelevant? What is the source of our knowledge of these axioms? What is the epistemological relationship between arithmetical laws and the particularities of numbers?

Is Theism More Rational Than Agnosticism: A Critique of Arguments for the Necessary Existence of God?

by David Kimweli

This paper engages the controversial question: is agnosticism a more rational opinion than theism? The paper examines the primary arguments for the necessary existence of God where Kant left it; having refuted the ontological, first cause, and design proofs and putting forth the necessity of God for the possibility of moral experience. After detailing Kant's view of transcendental morality, I then counter this view with the instrumentalist argument, first made by John Dewey, that sound moral judgments are made employing the same methods we can apply to any experience—the adaptive need to transform our environment beneficially. I make the case that Dewey's instrumentalist moral theory is superior to Kant's transcendental one, as it provides a simpler and more scientific rationale for moral experience. Lastly, I make the case that while belief in God has the potential to influence believers to live morally and is thus in a Deweyan sense instrumental, it has no factual basis and no moral or logical necessity, and as a result, an uninformed and irrational alternative to skepticism.

Rules and Realism: Remarks on the Poverty of Brute Facts

by J. Jeremy Wisniewski

In this paper, I offer a critical reconstruction of John Searle's argument for what he calls 'External Realism.' I argue that Searle's thesis is in fact ambiguous, and hence that it cannot establish the existence of brute entities (even if it can establish that we must presuppose an external world). I further argue that, once properly understood, constitutive rules can be shown to be prior to, rather than dependent on, what Searle calls 'brute facts' — and hence that Searle's analysis reverses the order of priority between rules and brute facts.

What is a Value Judgement?

by Georg Spielthener

The purpose of this paper is to clarify the concept of a value judgement. I present here my view on this problem, which is a version of non-descriptivism (or non-cognitivism) that is similar to but not identical to traditional non-descriptivist theories. The thesis I want to explain and argue for is that *S* makes a value judgement about *x* if and only if *S* expresses his attitude towards *x*. I explain first explain this thesis by (I) clarifying the concept of an attitude, in (II) I defend the identity between having an attitude towards something and evaluating it, in (III) I distinguish value judgements from judgements that only seem to be evaluative, in (IV) I clarify what I mean by ‘expressing an attitude’, and in section (V) I give a concise argument for my view.

Hyper Libertarianism and Moral Luck

by Gerald K. Harrison

This paper argues that if the principle of alternate possibilities is false, as many now believe, then there is a non-question begging reason to favour a hyper libertarian position over compatibilism. It will be argued that only a hyper libertarian position has the resources to provide a principled explanation of the reality of moral luck, something a compatibilist position cannot do.

WITTGENSTEIN, CONSCIOUSNESS, AND THE MIND

Pär Sundström

In traditional accounts of mental phenomena, consciousness often played a central role. For example, Locke, at one point in the *Essay*, considers whether it might be the case that we always think but are not always conscious of it. He dismisses this idea on the ground that, to think *is* to be conscious that one thinks: «'tis altogether as intelligible to say, that a body is extended without parts, as that any thing *thinks without being conscious of it*, or perceiving, that it does so. ... thinking consists in being conscious that one thinks» (book 2, chap. 1, §19).

In more modern accounts of mental phenomena, however, the role of consciousness is often downplayed. Instead, the mind is often accounted for in terms of «outer» or «objective» phenomena like mechanisms, dispositions, abilities and even environmental features. Different versions of this «outward turn» in the theorising about the mental can be found in the psychoanalytic literature, in behaviourism in philosophy and psychology, and in cognitive science.¹

In this paper, I examine one instance of this trend. In a series of passages, the later Wittgenstein criticised the idea that each mental phenomenon must involve some «inner state or process», such as an experience, sensation, feeling, or inner imagery. His treatments of this idea concerns a variety of specific mental phenomena, including *searching for a red flower* (*BB*, 3),² *comparing from memory* (*BB*, 85ff.), *understanding how to continue a series of numbers* (*PI*, §§151ff.), and *reading* (*PI*, §§156ff.). These passages are among the most well-received of the teachings of the later Wittgenstein. Sympathetic readers have cherished them, and critical readers have usually turned their attention elsewhere. According to many sympathisers, the passages show that one may exemplify any of the phenomena at issue without being in any particular type of conscious state.

I shall argue that the passages do not warrant this conclusion. Moreover, the conclusion is arguably false. For at least some of the phenomena in question, a strong case can be made for holding that in order to exemplify them, it is *necessary* that one is in a particular type of conscious state.

¹ It is true that there has been, in philosophy and also to some degree in cognitive science and the neurosciences, a renewed interest in consciousness lately. But this interest often coexists—at least in philosophy—with the view that consciousness is inessential for much or most of our mental life. That combination of attitudes is expressed by, e.g., Chalmers (1996). Despite being deeply fascinated by consciousness, Chalmers maintains that large domains of the mental can be accounted for without any invocation of it.

² I use the following abbreviations. *BB*: *The Blue and Brown Books*. *PI*: *Philosophical Investigations*. *Z*: *Zettel*. *R1-R2*: *Remarks on the Philosophy of Psychology*, vol. 1-2.

This is not to embrace the view—which might have been Locke’s—that all mental phenomena are conscious. But it is to suggest that Locke and other thinkers of the past may have been *more* right about the importance of consciousness for mental phenomena than we are currently inclined to think. Even if the conscious mind does not exhaust the mind, it may be a rather large and important part of it.³

The remaining agenda is as follows. In section I, I present the relevant passages from Wittgenstein. In section II, I document some conclusions that have been drawn on the basis of these passages. In section III, I assess what the passages show and do not show.

I Wittgenstein’s passages

In *The Brown Book*, Wittgenstein considers a scenario where an agent *A* shows a colour sample to another agent *B* upon which *B* goes and fetches an object that has the same colour as the sample, using his memory of the sample. In this case, we may say that *B* compares the colour of the object he sees with the colour of the sample he has just seen. But what makes it the case that *B* compares the two objects? What, for example, distinguishes the case where *B* compares the two objects from the case where he just by chance picks up an object that has the same colour as an object he has recently seen?

According to Wittgenstein, we are tempted to think that an essential component of such comparing is «a specific experience of comparing and recognising» (*BB*, 86); unless the agent *B* has such an experience, his performance is not one of comparing.

However natural this thought is, Wittgenstein believes it is wrongheaded. To convince us of this, he urges us to examine «closely» what really unites performances that we are prepared to label ‘comparing’. If we perform this close inspection, we find no *one* type of experience characteristic of comparing. Instead, we find that there is a great number of states of mind, all «*more or less* characteristic of the act of comparing». They include:

memory images, feelings of tension and relaxation, satisfaction and dissatisfaction, the various feelings of strain in and around our eyes accompanying prolonged gazing at the same object, and all possible combinations of these and many other experiences (*BB*, 86).

Individual acts of comparing are thus, according to Wittgenstein, a diverse lot. They resemble and differ from one another in various ways, but there is no one feature common to all of them.

We find that what connects all the cases of comparing is a vast number of overlapping similarities, and as soon as we see this, we feel no longer compelled to say that there must be some one feature common to them all (*BB*, 87).

Dialectics reminiscent of this passage show up in many other places. I have already mentioned the discussions of *searching for a red flower*, *understanding how to continue a series of numbers* and *reading*. In yet other passages Wittgenstein discusses what is involved in *intending a picture to be of so-and-so* (*BB*, 32f.), *believing what one says* (*BB*, 144ff.), *counting a number of objects* (*BB*, 149f.), *recognising* (*BB*, 165f.), *pointing to an object’s*

³ I am not alone in suggesting this. Notably, John Searle (1992) has argued vigorously that the role of consciousness in the mind has been underestimated in recent philosophy of mind and cognitive science. Galen Strawson (1994, chapters 6 and 11) also argues this point. I view my undertaking in this paper as supplementing these arguments. While Searle and Strawson defend the central role of consciousness in the mind mainly against claims arising out of cognitive science, I defend the same general view against claims arising out of the Wittgensteinian tradition.

shape (PI, §§33ff.), *intending* (PI, §591), *saying 'it'll stop soon' meaning the pain* (PI, §§666ff.), *writing a letter to so-and-so* (Z, §7), *looking for a photograph in a drawer* (Z, §8), *saying 'come here!' meaning a person A* (Z, §21f.), and *lying* (Z, §189f.). In each case, he considers the relation between the phenomenon under discussion and some inner state or process—such as an experience, sensation, feeling, or inner imagery—or class of such. And in each case, he suggests that no state or process out of the range considered is essential to the phenomenon.

A note about the relation between these passages: As has been observed (see, e.g. Malcolm, 9), it seems natural to understand a subset of the passages as different inquiries into one more general phenomenon, namely the phenomenon of *meaning* something by, say, a gesture or expression. This subset may include (all or some of) the passages about *intending a picture to be a portrait of so-and-so*, *believing what one says*, *counting a number of objects*, *pointing to an object's shape*, *saying 'it'll stop soon' meaning the pain*, *writing a letter to so-and-so* and *saying 'come here!' meaning a person A*. It is less natural, however, to understand *all* the passages as different inquiries into the general phenomenon of meaning something by a gesture or expression. For example, it is unnatural to construe the passages concerning *comparing*, *searching* and *recognizing* as inquiries into this general phenomenon.

I mention this to caution against quick generalisations over the class of mental phenomena investigated in Wittgenstein's passages. Even if one should become convinced that something is true about the general phenomenon of meaning something by an expression or gesture, the insight may not readily generalise to all of the mental phenomena under discussion. I will return, below, to the importance of recognizing the diversity of the phenomena under discussion.⁴

II Suggested lessons of the passages

What do the passages that we have just looked at show? According to Norman Malcolm, they have dramatic implications for how we should understand mental phenomena. Malcolm thinks we should conclude, on the basis of the passages, that one may exemplify any mental phenomenon regardless of «what goes on in one's mind or thoughts»:

We are tempted to think that your meaning the color must have been something that went on in your mind. But it might be that what went on in your mind or thoughts had *nothing* to do with what you meant. ...Instead of looking *inside* ourselves we should be looking *around* us, at the context in which our words and pointing are located. We should be searching horizontally instead of vertically. This temptation to look in the wrong direction besets us whenever we are perplexed about the concepts of mind. Wittgenstein's admonition applies to all of them (1970, 15-6).

Malcolm's suggestion is perhaps unusually radical. But with a few reservations, it is representative of what many readers take Wittgenstein to have shown in the passages under consideration.⁵ The reservations are three: First, other commentators are usually more guarded than Malcolm in their claims about the range of phenomena to which Wittgenstein's results apply. They usually don't draw conclusions about mental phenomena generally, but only about some subset of them. Exactly which subset they have in mind is not always explicit, but my

⁴ See section 3, claim (4).

⁵ In the following, I am trying to summarise assessments such as those arrived at by Budd, pp. 21-6, McGinn, pp. 93-117, Pitcher, chapter 11, Kripke, pp. 40-51, and Baker and Hacker, pp. 358-9.

impression is that many readers take Wittgenstein to have shown something significant about at least the range of phenomena considered in the passages listed in section I above.⁶ Second, not all commentators formulate Wittgenstein's result in terms of what goes on in all of «mind or thought». Many formulate them instead, more cautiously, in terms of what goes on in «consciousness», or what goes on that can be «introspected».⁷ And third, not all commentators take Wittgenstein to have shown that exemplifying a given phenomenon may have «*nothing to do*» with what goes on in consciousness. Mostly, commentators have concluded that some relation fails to hold between exemplifying a given phenomenon and being in a particular *type* of conscious state. In some places, it is concluded that there is no type of conscious state the having of which *amounts to* exemplifying the phenomenon in question.⁸ In other places, it is concluded that there is no type of conscious state the having of which is *necessary* to exemplifying the phenomenon.⁹ And in yet other places, the conclusion is that there is no type of conscious state that a subject is *always* in when he or she exemplifies the phenomenon.¹⁰

To summarise this, each of the following conclusions have been drawn, by more than one commentator, on the basis of Wittgenstein's passages:

(CI) There is no phenomenon P (out of the range discussed in the passages listed in section I) and type of conscious state C, such that: being in C *is* to exemplify P.

(CN) There is no phenomenon P (out of the range discussed in the passages listed in section I) and type of conscious state C, such that it is *necessary* that: if one exemplifies P then one is in C.

(CA) There is no phenomenon P (out of the range discussed in the passages listed in section I) and type of conscious state C, such that it is *always* the case that: if one exemplifies P then one is in C.

On a natural understanding of these claims, they can be ordered with respect to strength, (CA) being the strongest and (CI) the weakest. That is, (CA) implies (CN) but not vice versa, and (CA) and (CN) both individually imply (CI), but not vice versa.

To simplify my discussion, I shall set the weakest of the claims, (CI), to one side. I'm not quite sure what expositors of Wittgenstein have meant by it, and a discussion of this would take us too far afield.

⁶ The impression is based on the following. While many commentators (for example, Pitcher, McGinn and Kripke) have had their focus on the phenomena of *meaning* and *understanding*, at least some (for example, Budd, and Baker and Hacker) have explicitly taken Wittgenstein to demonstrate something about a broader range of phenomena. And no commentator that I am aware of has discussed the possibility that Wittgenstein's results may apply to only some of the phenomena considered in the relevant passages.

⁷ See Kripke, pp. 49f., McGinn, pp. 104ff., Baker and Hacker, p. 359, and also Budd, p. 25.

⁸ See Kripke, pp. 41f., McGinn, p. 96, and Pitcher, pp. 259f.

⁹ See McGinn, pp. 96-7, Pitcher, pp. 259f., Baker and Hacker, p. 359, and Budd, p. 26.

¹⁰ See Pitcher, p. 260, and Kripke, pp. 43ff.

Setting (CI) aside, I shall in the following cast doubt on (CN) and (CA). I think it is doubtful that these claims are true. And I think it is *clear* that they are not justified by the passages in question. I turn to discuss these issues now. Since casting doubt on the (weaker) claim (CN) amounts to casting doubt on the (stronger) claim (CA), I will target the former.

III What the passages show and do not show

To assess what Wittgenstein's passages show and do not show, it is crucial to distinguish two different interpretations of the temptation of thought that Wittgenstein identifies and tries to exorcise.

On the first interpretation, the temptation is to say, for any phenomenon P under consideration, that to exemplify P one must be in any one out of a *narrow* range of states of mind. I shall not try to make absolutely precise what this narrow range of states of mind is, but we shall think of it as including *sensations*, *mental images*, and *feelings*, while not including *awareness that* something is the case. Thus, on this interpretation, the temptation is to say things like: to exemplify P one must have a particular type of *sensation*, or *feeling*, or *image*; but the temptation is not to say that: to exemplify P one must be *aware that* something is the case. I will sometimes refer to this as the «sensationalist» interpretation of the temptation.

On the second interpretation, the temptation is to say, for any phenomenon P under discussion, that to exemplify P one must be in any one out of a *broader* range of states of mind. On this interpretation, the temptation *is* (in part) to say that: to exemplify P one must be *aware that* something is the case.^{11,12}

With this distinction in place, I shall in the remainder of the paper try to make the following claims plausible:

(1) If Wittgenstein's temptation is understood narrowly or sensationalistically, it is right to reject it in the case of each phenomenon under discussion (comparing, reading, understanding, and so forth). I can compare from memory, or exemplify any of the other phenomena under discussion, without having any particular type of feeling, sensation or mental image.

(2) However, rejecting the temptation as narrowly or sensationalistically conceived does not by itself give us any reason to conclude that there is no type of *conscious* state that one must be in to exemplify these phenomena. This is because there are arguably conscious states other than sensations, feelings, and imagings; and consequently, there may well be some type

¹¹ Distinctions that *seem* similar to the one I am drawing here have been recognised as important to the assessment of the passages under discussion. Examples are Budd's distinction between an «extrinsic» and an «intrinsic» conception of states of consciousness (Budd, pp. 22-3), and McGinn's distinction between «quotidian» and «queer» conscious contents (McGinn, pp. 7-9). (See also Kripke, pp. 51ff., and *PI*, §§187-97.)

Despite drawing these distinctions, however, neither Budd nor McGinn pays any serious attention to the second—broad—interpretation of the temptation above. This may suggest that their distinctions do not, after all, coincide with mine. (*If* they have drawn just the distinction that I draw, then I believe they have failed to appreciate the consequences of this distinction for the assessment of Wittgenstein's passages.)

¹² Exegetical note: The two interpretations do *not* seem to me to do equal justice to *each* of the passages under consideration. I believe some passages suggest a narrow interpretation of the temptation that they discuss, while other passages suggest a broad interpretation. (And in yet other passages, there are aspects that suggest a narrow interpretation and other aspects that suggest a broad interpretation.) But whether this is right or not will not matter to my discussion.

of *conscious* state I need to be in, in order to compare, even though there is no sensation, feeling or image I need to have in order to do so.

(3) We *would* be justified in thinking that no conscious state is necessary for exemplifying any of the phenomena in question *if* we could reject a certain version of the temptation as broadly conceived.

(4) However, it is doubtful that this temptation *should* be rejected in the case of each phenomenon. Certainly, Wittgenstein's passages don't provide sufficient reason for such a sweeping rejection.

While I shall say something in favour of each of these claims, it should be obvious that I have a stake only in (2) and (4), and I will consequently devote most of the space to these. I happen to believe that (1) and (3) are true as well; but should they be false, that would not harm my argument. I will attend to the claims in turn.

(1) *We should reject Wittgenstein's temptation as narrowly, or sensationistically, conceived.*

I will not spend much energy defending this claim. The claim is, in effect, that there is no particular type of sensation, image or feeling that I need to have in order to compare, or exemplify any other of the phenomena at issue. And this is clearly right. It may not have been *obvious* or *widely appreciated* before Wittgenstein pointed it out to us. Thus, philosophers like Hume and Locke probably thought otherwise. But by now, philosophers have largely absorbed this particular lesson of Wittgenstein's.

(2) *But to reject the temptation as narrowly or sensationistically conceived does not by itself give us any reason to think that there is no type of conscious state that one must be in to exemplify any of the phenomena at issue (because there are arguably conscious states other than sensations, feelings, and imagings).*

This claim requires much more defence. To defend it, I introduce the following claim:

(AN) It is necessary that: if one compares then one is *focally action aware* that one is comparing.

Later on, I shall defend this claim.¹³ But for present purposes, much less is required. First, I shall explain what it is to be «focally action aware» that one is doing so-and-so. Second, I shall try to make plausible that such focal action awareness is a type of *conscious state*. If this is so, then there are conscious states other than sensations, feelings, and imagings. And then, there may well be some type of conscious state—for example, the state of being focally action aware that I am comparing—that I need to be in, in order to compare, even though there is no type of sensation, feeling or image that I need to have to do so. And that suffices to establish (2).

What, then, is it to be *focally action aware* that one is doing so-and-so? It will take some work and care to pinpoint this phenomenon. But as a first characterisation, I am typically focally action aware that I am doing so-and-so when I'm in a position to report, without *special prompting*, that I am doing so-and-so. By 'special prompting', I have in mind here the act, which can be performed by myself or somebody else, of bringing to my attention the

¹³ See the argument for claim (4) below.

possibility that I am doing so-and-so. Such prompting might take the form of the question ‘Are you (or: am I) doing so-and-so?’.

To illustrate, suppose I am comparing. And suppose further that, as I do so, there is a set of propositions that I’m in a position to sincerely express, *without* special prompting, in reply to the question ‘What are you now doing?’ and another set of propositions that I’m in a position to express in reply to this question only *given* prompting. Let us suppose that the first set of propositions includes *I am comparing* and *I am helping so-and-so find something that agrees with this sample*, and that the second set includes *I am knitting my brows*, *I am exerting pressure on the soles of my shoes*, *I am trying to make up to so-and-so for letting him down the other day*, and *I am suppressing an urge to do something more fun*. In such a case, I am typically focally action aware that I am comparing and that I am helping so-and-so find something that agrees with the sample. In contrast, I am not focally action aware that I am knitting my brows or that I am trying to make up to so-and-so for letting him down the other day—although I may in *some* sense be aware that I am doing these things.

I intend this characterisation to *fix the reference* of ‘focal action awareness’ rather than defining the term. Thus, focal action awareness is not *defined* as some kind of behavioural disposition. It is, rather, *ostended* as that kind of awareness that is characteristically present in the kind of circumstances described and exemplified above.

A few words need to be said about the relation between focal action awareness and *attention*. Focal action awareness does not require attention. Or at least, being focally action aware *that one is doing so-and-so* doesn’t require that one attends to the *fact that one is doing so-and-so*. Just think about a typical case where you are comparing, say, the colours of two objects and are in position to report, without special prompting, that you are doing so. In the typical such case, what you are attending to is surely *the colours of the two objects*. You are not attending to the fact that you are now comparing them. (And, as is often pointed out, if you start to attend to the fact that you are comparing them, this will distract you from what you are doing and make it harder for you to tell whether the objects match or not.) But while you are *attending* to the colours of the two objects, you still have a certain kind of *awareness* of the fact that you are comparing them. If you are asked what you are doing, you can effortlessly report, without special prompting, that you are comparing. And this is plausibly *because* you have this awareness—«focal action awareness»—of what you are doing.

That you are focally action aware that you are doing so-and-so does not mean, then, that the fact that you are doing so-and-so is the focus of your *attention*. It means, rather, that it is in the focus of your *action awareness*, that is, your awareness of what you are *doing*. When you compare two objects, you are typically aware, in some sense or other, of doing a number of things. And, even if you attend only to objects outside of you and not to any of the things you are doing, it is often or always possible to distinguish what you are more *focally* aware of doing from what you are less focally aware of doing. If you are asked what you are doing, there are some true answers you can provide *without* special prompting, and other true answers you can give only *given* special prompting. The former answers typically express what you

are more focally aware of doing; the latter answers what you are less focally aware of doing.¹⁴

Now, some may doubt that focal action awareness is present in the kind of cases I have described. Thus, it is sometimes suggested—often on phenomenological grounds—that, when all your attention is directed «outward», you are in fact not in any way aware that you are comparing *even if* you can report that you are doing so without special prompting. Your ability to effortlessly *report* on what you are doing may foster an illusion that you have such awareness. But closer phenomenological inspection will tell you that, in reality, you *become* aware that you are comparing only when you are asked what you are doing.¹⁵

However, I don't see any reason—whether phenomenological or of some other kind—to think that, in this kind of case, I become aware that I am comparing only when I'm asked what I am doing. Let's distinguish two kinds of cases in which you are comparing the colours of two objects and are asked, in the midst of this performance, what you are now doing. In one kind of case, this question causes you to shift some of your attention away from the colours of the objects and towards the fact that you are comparing them. In the second kind of case, the question doesn't cause any such shift in attention. (Think, for example, about situations where you are, as we say, very «focussed on your task». In those situations, you may report that you are comparing when asked what you are doing, but while reporting this you remain just as attentive to the colours of the object. In such cases, your report will often be somewhat absent-minded). I trust that both cases are familiar. Now, I don't find, in *either* of these cases, any evidence that I become aware that I am comparing only when I'm asked what I am doing. Consider the latter case first. In this case, there are indeed some changes in my awareness as the question is asked. First, I become aware that a question is being asked, and later, I am aware that I am answering this question. But I fail to detect any changes in my awareness in addition to those. In particular, I find no evidence that I suddenly come to realise that I'm comparing only when the question is posed. Rather, it seems to me that I'm reporting something I was aware of all along. This phenomenological impression is to some degree confirmed by the very fact that I can effortlessly report what I'm doing while staying fully «focussed on the task». It is relatively easy to report something that one is already aware of. If I became aware that I'm comparing only when asked what I am doing, this would be cognitively more disruptive, and it would be at least slightly more surprising that I could stay fully focussed on my task while having this realisation and then reporting it. Consider now the former kind of case, where the question what I'm doing causes a shift in my awareness. In this case, there is a change in my awareness in addition to my being aware of a question and my answer to it, because there is a change of attention. When I'm asked what I'm doing, I come to attend (or attend more) to the fact that I am comparing. But, again, I see no reason to think that I only then *become* aware of the fact that I'm comparing. As far as I can tell from any phenomenological inspection, the truth is rather that I come to attend (or attend more) to a fact that I was *already* aware of but didn't pay (much) attention to.

¹⁴ A parallel distinction between the focus of attention and the focus of some mode-specific awareness can be made in other cases. Thus, suppose I attend to something that I see. It may still be possible to distinguish more or less focal elements of my audition. If you were to ask me what I hear, I might be able to answer that I hear a tune *without* special prompting, but answer that I hear the faint noise of distant traffic only *given* special prompting.

¹⁵ This claim is made by Wakefield and Dreyfus, 268.

Now, focal action awareness that one is doing so-and-so is, I submit, a type of conscious state. Admittedly, this suggestion brings us into a disputed, and rather murky, territory. As has often been remarked, we possess many concepts of consciousness.¹⁶ And it is true that there is at least *one* sense of consciousness such that *some* philosophers deny that there are, in that sense, conscious states other than sensations, images, and feelings. Thus, Michael Tye has argued that sensations, imagings, and feelings are the only states that are «phenomenally» conscious; that is, they are the only states that are ever *like* something to be in.¹⁷ However, this view is controversial. Many—perhaps even most—philosophers of mind hold that states other than these—for example thoughts and desires—can also be like something to have and thus be phenomenally conscious.¹⁸

The issue is elusive. All parties of the dispute agree that there is something it is like for me *when* I, say, think a certain thought. Tye insists that what it is like is always solely a matter of what sensations, images, and feelings *accompany* the thoughts. But others insist that this is not so; that the thought *itself* can be like something to have. It is unclear, at least to me, what could settle this dispute.

In the present context, however, there is something more to say. As we are discussing what Wittgenstein's passages show and do not show, it is of interest to consider *Wittgenstein's* favoured criteria for a state's being conscious. And it seems that focal action awareness that one is doing so-and-so satisfies these criteria, at least, for being a conscious state.

When Wittgenstein tried to determine whether some psychological phenomenon was a «state of consciousness»—or even, sometimes, whether it was a *mental* state—he characteristically considered whether the phenomenon displayed a certain kind of duration, which he called «genuine duration». For example, in *Remarks on the Philosophy of Psychology*, he says:

I want to talk about a «state of consciousness», and to use this expression to refer to the seeing of a certain picture, the hearing of a tone, a sensation of pain or of taste, etc. I want to say that believing, understanding, knowing, intending, and others, are not states of consciousness. If for the moment I call these latter «dispositions», then an important difference between dispositions and states of consciousness consists in the fact that a disposition is not interrupted by a break in consciousness or a shift in attention. (And that is of course not a causal remark.) Really one hardly ever says that one has believed or understood something «uninterruptedly» since yesterday (*R2*, §45).

Think of this language-game: Determine how long an impression lasts by means of a stop-watch. The duration of knowledge, ability, understanding, could not be determined in this way (*R2*, §51.)

The general differentiation of all states of consciousness from dispositions seems to me to be that one cannot ascertain by spot-check whether they are still going on (*R2*, §57).¹⁹

¹⁶ See, e.g., Güzeldere, pp. 8-9, and Chalmers, pp. 25-31.

¹⁷ Tye, section 2. Familiarly, the specification of 'consciousness' in terms of 'what it is like' became widespread in philosophy through Nagel (1974).

¹⁸ Galen Strawson is the most emphatic and articulate defender of this view; see sections 1.3-1.4. But the view is also embraced by, e.g., Block (p. 230), Chalmers (pp. 9-10), and Flanagan (p. 64).

¹⁹ See also *PI*, §148, and p. 59, Insert (a); *Z*, §§72, 75-8, 81-5, 472; *R1*, §836; *R2*, §63.

It seems that focal action awareness that one is doing so-and-so satisfies the conditions for being a conscious state that one can extract from passages such as these:²⁰ To begin with, focal action awareness that one is doing so-and-so is interrupted by a break of consciousness. For example, if I am focally action aware that I am doing something and then suddenly fall into dreamless sleep, my focal action awareness is interrupted. Further, it makes sense to say that one has been focally action aware of something uninterruptedly for a certain period. For example, I can sensibly say that for the past half-hour, I have without interruption been focally action aware that I am working on this paper. Again, it is possible to determine by means of a stopwatch for how long someone is focally action aware of something. Thus, the kind of assessment I just made—that for half an hour I have been focally action aware that I am working on this paper—could have been made with greater accuracy by using a stopwatch. (To be sure, it will not always be possible to make this kind of assessment very precise even with a stopwatch, since the beginnings and end points of focal action awareness are not always sharply delimited. But in this regard, focal action awareness does not differ from sensations of pain or taste, which are cited by Wittgenstein as examples of states of consciousness.) Finally, one can ascertain by spot-check whether one is still focally action aware of something. For example, ten minutes ago I was focally action aware that I was working on this paper, and I can now determine that I am still focally action aware that I am doing that. In contrast, an hour ago I was focally action aware that I was brushing my teeth, and I can now determine that I am no longer focally action aware of doing that.

True, there is one condition for being a conscious state which Wittgenstein mentions and which is not satisfied by focal action awareness: focal action awareness is not interrupted by any shift in attention (for this condition, see the first of three passages quoted above). I may shift my attention, say, from substantive to formal aspects of the paper, and yet remain focally action aware that I am working on it. But then, it seems that this condition should at least not be a *necessary* condition for being a state of consciousness, by Wittgenstein's own lights. Wittgenstein thinks that hearing a tone and a sensation of pain are conscious states (first passage quoted above), and neither are those interrupted by any change of attention. If I hear a tone or feel a pain, I may shift my attention in various ways and still hear the tone or feel the pain.

So, focal action awareness that one is doing so-and-so is, arguably, a kind of conscious state. That is at least what many—perhaps most—contemporary philosophers of mind would think or allow. And it is also the verdict one seems to reach if one considers Wittgenstein's criteria for a state's being conscious. But if focal action awareness is a conscious state, then there are conscious states other than sensations, feelings, and imagings. And hence, there may well be a type of conscious state—for example, a certain state of focal action awareness—that I need to be in, in order to compare, even though there is no sensation, feeling or image that I need to have to do so. And that suffices to establish (2).

(3) We would be justified in thinking that no conscious state is necessary for exemplifying any of the phenomena in question if we could reject a certain version of the temptation as broadly conceived.

²⁰ It is not clear to me whether the conditions stated in the passages should be understood as necessary or sufficient or both. I only claim that whether understood as necessary or sufficient or both, focal action awareness seems to satisfy them.

To repeat: On the broad interpretation, the temptation is to say, for any phenomenon P under discussion, that to exemplify P one must be in any one out of a *broader* range of states of mind. On this interpretation, as opposed to the former interpretation, the temptation *is* (in part) to say that: to exemplify P one must be *aware that* something is the case.

Wittgenstein's discussion of reading—or one strand of this discussion at any rate—affords us an example of a temptation thus broadly conceived. Part of the formulation of the temptation in this case is: «A man surely knows whether he is reading or only pretending to read!» (*PI*, §159). Judging by this particular formulation, it seems that the temptation under discussion is to say that some sort of *awareness* or *knowledge* that one is reading is necessary for reading.²¹

If we could show this temptation to be false, then I think we could reasonably conclude that there is no type of conscious state that one must be in, in order to read; and analogous conclusions would be warranted for the other phenomena under discussion: for example, if awareness that one is comparing is not necessary for comparing, then it is reasonable to conclude that no conscious state is necessary for comparing.

Why would these conclusions be reasonable? After all, awareness that one is doing so-and-so is at most *one* type of conscious state. How can it be reasonable to conclude that *no* conscious state necessarily accompanies doing so-and-so on the basis of ruling out that *that* type of conscious state necessarily accompanies doing so-and-so?

The answer is that there is, arguably, *no more plausible candidate*, among the states of consciousness, for being a necessary accompaniment of comparing than the (focal action) awareness that one is comparing. Contrast the focal action awareness that one is comparing with, say, the feelings of tension and strain that Wittgenstein considers in *The Brown Book*. Clearly, focal action awareness that one is comparing is a *more* plausible candidate for being a necessary accompaniment of comparing than are feelings of strain and tension. Arguably however, there is no more plausible candidate, among the states of consciousness, for being a necessary accompaniment of comparing than the focal action awareness that one is comparing. That's why it would be reasonable to conclude that *no* conscious state is necessary for comparing, if we could rule out that focal action awareness that one is comparing is not necessary for doing so.²²

As I said above, I think this is correct, but it is not essential to my argument. Essential, on the other hand, is the following point:

²¹ Other passages that invite a broad interpretation of the temptation under treatment are the discussions of writing a letter to so-and-so, and of looking for a photograph in a drawer, in *Z*, §§7 and 8 respectively.

²² To draw this kind of conclusion, one must be careful to really identify the *most* plausible candidates, among the states of consciousness, for being necessary accompaniments of whatever phenomena one investigates. An illustrative failure to do so is provided by one passage from Malcolm. In the relevant passage (p. 10), Malcolm initially follows Wittgenstein in rejecting the idea that a person who points to an object's colour must attend to, and in that sense be aware of, *the object's colour*. With this rejection, I have no quarrel. But on the basis of it, Malcolm proceeds to the conclusion quoted above, that pointing to an object's colour may have «nothing to do» with what goes on in mind or thoughts. And this conclusion is not justified. It is not justified because attending to, or being aware of, an object's colour is *not* the most plausible candidate, among the states and processes of mind and thought, for being necessary to pointing to an object's colour. A more plausible candidate is being aware that *one is pointing to* the object's colour. (And clearly I can be aware that *I am pointing to* an object's colour without attending to, or being aware of *the object's colour*.)

(4) *It is, however, doubtful that this version of the broad temptation should be rejected in the case of each phenomenon under consideration. Certainly, Wittgenstein's passages don't provide sufficient reason for such a general rejection.*

To support this, I return to the claim (AN), which says that in order to compare, one must be focally action aware that one is comparing. If (AN) is true, then a certain version of Wittgenstein's temptation (broadly conceived) is true for at least *one* of the phenomena under discussion, namely *comparing from memory*. Above, I explained what it is to be focally aware that one is comparing, but I did nothing to defend the claim (AN). I will defend it now.

As a preliminary to defending (AN), I want to make two points. The points are both rather obvious, but they have nevertheless been blurred in some commentaries on Wittgenstein. Taking notice of them will clear the way for a sober assessment of (AN).

First point: Comparing is not an *ability* or *capacity*. To be sure, comparing presupposes at least one ability, namely the ability to compare. But it *is* not that ability, or any other ability. Comparing is rather the *exercise* of the ability to compare. Now, it is evident that I may be *able* to compare without being focally action aware that I am comparing (and also without being aware that I am able compare). But this must not be confused with the possibility of *comparing* without being focally action aware that one is comparing.²³

Second point: The question of what is necessary for *falling* under a concept (say, the concept *comparing*) must be distinguished from the question of what is necessary for *understanding* the concept, as well as from the question of what is necessary for *coming* to understand the concept. To understand, or come to understand, the concept *the winner of the race* one doesn't have to be the first one to cross the line. In fact, one doesn't have to be in the race at all. But to *fall* under this concept—that is, to win the race—one must be the first one to cross the line. Now, it's clear that one can *understand* the concept of comparing without being focally action aware that one is comparing (and also without being aware that one understands the concept). And it seems plausible, if perhaps less obvious, that one can *come* to understand this concept without being or ever having been focally action aware that one is comparing (and also without being aware that one is coming to understand the concept). But

²³ The distinction between abilities and exercises of abilities is often blurred in commentaries on Wittgenstein. For example, Budd claims that Wittgenstein, in the *Investigations*, by 'reading' understands «the *ability* to follow certain kinds of rule» (p. 27; emphasis added). But what Wittgenstein says is: «reading is here the *activity* of rendering out loud what is written or printed» (*PI*, §156; emphasis added.) Later on, Budd infers that the rules of chess need not be present in the mind of someone who *decides* to play chess, on the ground that those rules need not be present in the mind of someone who *can* play chess (p. 37).

It is worth making a note about *understanding* in this context. The word 'understanding' can be applied to (i) something one can be said to do on a given occasion—such as understanding (or meaning) something by an expression or gesture—as well as to (ii) something one cannot be said to do on a given occasion—such as understanding (or knowing) the meaning of an expression. (By saying that understanding in the former sense is something one can be said to *do* on an occasion, I don't mean to suggest that it is necessarily an *action*. The main point is that it is something that takes place *at a time*. Thus, I can understand—or mean—one thing by 'bank' on one occasion, and another thing by 'bank' on a different occasion. But one may object to saying that these are *actions* on the ground that it seems nonsensical to *order* somebody to understand or mean an expression in a certain way. Cf. *Z*, §51.) Now, understanding in sense (ii) may well be an ability; but understanding in sense (i) is not. To understand (or mean) something by an expression on a given occasion is the *exercise* of an ability. Nevertheless, the distinction between these two kinds of understanding has been blurred in commentaries on Wittgenstein as well.

none of this implies that one can *compare* without being focally action aware that one is comparing.²⁴

With these potential confusions out of the way, I shall try to bring out the case for holding (AN). My strategy will be to contrast comparing with some other things one can do.

In the case of many things one can do, it is indeed clear that one can do them without being focally action aware that one is doing them. Breaking a window and touching wet paint are cases in point. I can obviously break a window or touch wet paint without being focally action aware that I am doing any of this, and indeed, without being focally action aware of anything; I can do these things when I am in dreamless sleep. The same is true of performances often associated with comparing, such as *picking* something that in fact agrees in some respect with a given sample.

In the case of some other things one can do, it may be necessary that one is focally action aware of *something*, and there may even be some necessary *restrictions* on what one is focally action aware of when one does them. Cases in point may be displaying generosity and proceeding brutally. But it is again clear that I can do either of these things without being focally action aware that I am doing precisely *it*. Focal action awareness that one is displaying generosity is clearly not necessary for displaying generosity, for example.

In contrast, it is less clear that I can *compare* something with something else without being focally action aware that I am doing so. Imagine a case: A person S is standing by a bookcase. His or her eyes are running over the shelves. After a while, he or she grabs a book. But suppose it is not true that S is in a position to report, without special prompting, that he or she is comparing. We may suppose that S might without special prompting report on doing other things (like *I'm waiting for so-and-so* or *I'm trying to figure out what so-and-so's reading habits are*). Or we may suppose that there is nothing that S can report on doing without special prompting. In either case, we have reason to think that S is not focally action aware that he or she is comparing. And it seems that we thereby have strong reason to withdraw or contest any judgment to the effect that S *is* comparing. This is true if we conceive of S from a third-person perspective, but also if we do so from a first-person perspective: If *I* in this kind of situation am not focally action aware that I am comparing, it seems that *I* thereby have strong reason to contest any judgment to the effect that I am comparing. Even if there are features of the situation that indicate that I am comparing, the mere fact that I am not focally action aware of doing so gives me reason to think I am not. The contrast with the case of (say) *picking* something that in fact matches a sample is striking: even if I lack focal action awareness that I am picking something that matches a given sample, I can very easily be convinced by other evidence that I am doing so.

I take this to show that (AN) has, at the very least, some initial plausibility: if the absence of focal action awareness of comparing gives us reason, both from the first-person and third-

²⁴ These distinctions have been blurred in commentaries on Wittgenstein as well. For example, Baker and Hacker say this: Since there is no such thing as private samples ... psychological concepts cannot be explained by introspection. All introspective reports can provide is a variety of experiences or phenomena accompanying understanding, believing, fearing, wanting. These accompaniments are neither necessary nor sufficient for the application of such psychological predicates (p. 359.)

This passage says, first, that a concept like fearing cannot be *explained* by introspection, and second, that nothing reportable on the basis of introspection is necessary or sufficient for the *application of predicates* like «fearing». But these claims, uttered in one breath, are very different. The latter is a claim about when somebody *falls* under the concept *fearing*; the former is a claim about what it takes to *acquire* it.

person perspective, to infer the absence of comparing, then we have reason to believe that there is some intimate connection between the two.

Are there any convincing reasons to reject (AN), despite its initial plausibility? I'm not aware of any. An exhaustive discussion of this issue would make this paper too long. I shall confine discussion to two challenges that can be derived from Wittgenstein's passages. I hope that my responses to these will contribute to make (AN) seem plausible.

The first challenge I adapt from Wittgenstein's discussion of reading, in the *Investigations*, §§156ff. Consider a hypothetical drug which had no impact on people's capacities for comparing, but which made them unable to be focally action aware of what they were doing. If such a drug is possible, then it is obviously possible to compare without being focally action aware that one is comparing.

However, the question is precisely whether such a drug *is* possible. After all, the connection between comparing and focal action awareness that one is comparing does not seem entirely accidental. Comparing is an activity that requires some degree of sustained attentional involvement. It seems to be—in every sense—impossible that somebody just *happens*, by chance, to compare a set of objects with a given sample. (Again, there is a striking contrast with the case of *picking* an object that in fact matches some sample: somebody may perfectly well happen, by chance, to do that.) Now, it may be that the sustained attentional involvement required for comparing in turn requires that one be focally action aware that one is comparing. Perhaps such awareness is needed in order to keep sufficient cognitive resources enlisted for the task at hand. I'm not insisting that this is true, but the idea does not seem ludicrous. And if it is right, then any drug that impaired an agent's capacity to stay focally action aware of what he or she is doing would *necessarily* impair the agent's capacity for comparing.

To be sure, we may sometimes feel that we are able to imagine or conceive of someone who—for example under the influence of a drug—compares without being focally action aware of doing so. But this may reflect that we are at those moments not too clear about what it is to compare and/or be focally action aware. We know from other cases that we are prone to make such mistakes: Someone who is sufficiently ignorant of geometry may find it conceivable that the square of the hypotenuse of a Euclidean right-angled triangle is not equal to the sum of the squares of its other two sides. Nevertheless, it is *necessary* that the square of the hypotenuse of a Euclidean right-angled triangle is equal to the sum of the squares of its other two sides.²⁵ Similarly in this case: If it seems to us conceivable that someone compares without being (focally action) aware of doing so, this may just reflect that we are unclear about what it is to compare and/or be (focally action) aware; if we had a sufficiently clear understanding of this, it would perhaps be apparent that it's necessary that anyone who compares is focally action aware that he or she is comparing.

The second challenge I want to consider is this: Every now and then, we forget what we are doing in the midst of an activity. After a while we may ask ourselves, 'What am I doing again?' and only then does awareness return: 'Oh yeah, I am comparing'. This may seem to show that it is not only *conceivable* that someone compares without being focally action aware that he or she is comparing, but that this happens in fact, and so is obviously possible.²⁶

²⁵ This is, of course, what Arnauld pointed out in the fourth set of objections to Descartes' *Meditations*.

²⁶ This objection echoes the short discussion of *looking for a photograph in a drawer* in Z, §8.

In reply to this challenge, I think it is unproved that one in fact ever continues to compare while being thus forgetful. There is of course a sense in which one may compare the whole day long even if, during the day, one takes both a lunch break, a nap and a dinner break. That is the sense of ‘comparing’ in which comparing does not have what Wittgenstein called ‘genuine duration’. But there is another sense in which comparing does have genuine duration. This is the sense in which I stop comparing when I take a break, and resume comparing after the break.²⁷ It should be obvious that we have been concerned with comparing—as well as searching, pointing, reading, and so forth—in this latter sense all along. And it is unproved that I ever continue to compare in *this* sense while I am forgetful of what I am doing.

To be sure, a maximally reflected verdict on (AN) would have to consider further objections. But I hope the above discussion has contributed to make (AN) seem plausible.

Now, (AN) is a claim only about *comparing*. What about the other phenomena under discussion, searching, pointing, understanding, and so forth: Can it be argued, along similar lines, that to exemplify any of these, one must be focally action aware that one is doing so?

At this point, generalisations cannot be made across the board. In the case of *some* of these other phenomena, this claim may be defensible. For example, it may be defensible to hold that in order to search for a red flower I must be focally action aware that I am searching for a red flower. But this kind of claim is clearly not plausible for *all* the phenomena at issue. For example, it is clearly false that I understand how to continue a series only if I’m focally action aware that I understand this. Similarly, I think it is clearly false that I intend to do so-and-so only if I’m focally action aware that I intend to do so-and-so.

Thus, we should recognise that Wittgenstein’s temptation as *broadly* conceived has very different degrees of plausibility depending on which phenomenon we are talking about. While the temptation is clearly indefensible in the case of, say, understanding and intending, a certain version of it may well be defensible for, say, comparing and searching.

IV Conclusion

I have distinguished two interpretations of the kind of temptation Wittgenstein tried to exorcise in his investigations into comparing, searching, understanding, reading, and other psychological phenomena. I have argued that (1) on a narrow or sensationalistic interpretation of the temptation, it is right to reject it in each case. But (2) the rejection of these narrow temptations does not give us reason to draw the sort of anti-Lockean conclusions about consciousness and the mind that many have taken the passages to establish. Specifically, it does not warrant the conclusion that no type of *conscious state* is necessary for, say, comparing or searching or reading. (3) We would, on the other hand, have reason to draw this conclusion if we could show that a certain version of the temptation as broadly conceived should be rejected. But then (4) it is far from clear that this temptation should be rejected in the case of each phenomenon under discussion. Certainly, Wittgenstein’s passages do not provide sufficient reason to do so.

Wittgenstein cautioned us, in the *Investigations*: «When we do philosophy, we should like to hypostatize feelings where there are none» (§598). This is no doubt a valuable warning. If we think that each psychological phenomenon essentially involves some feeling, then we

²⁷ Familiarly, Wittgenstein himself suggested that many psychological predicates display this ambiguity, and that our understanding of psychology would be enhanced if we had different words for the different senses. See *PI*, §577.

hypostatise feelings where there are none, and thus distort our understanding of the mind. Fortunately, philosophers have largely absorbed this lesson. But 50 years after the *Investigations*, it seems we have reason to issue a complementary caution: When we do philosophy, we are sometimes led to deny the existence of consciousness where there seem, after all, to be some. For example, we are sometimes led to conclude, on the basis of the claim that one can exemplify a given mental phenomenon without having any particular type of *sensation* or *image* or *feeling*, that one can exemplify that phenomenon without being in any particular type of *conscious state*. And then we end up denying consciousness where there (essentially, necessarily, or universally) is some, and *thus* distort our understanding of the mind. This lesson may be just as consequential for our understanding of the mind as the one that Wittgenstein taught. And I think it remains to be fully absorbed.²⁸

References

- Baker, Gordon and Hacker, Peter (1986) «The Grammar of Psychology: Wittgenstein's *Bemerkungen Über die Philosophie der Psychologie*», in *Ludwig Wittgenstein: Critical Assessments*, vol. 2, Stuart Shanker (ed.), London: Croom Helm, pp. 352-72. Originally published in *Language and Communication*, 2, 1982, pp. 222-44.
- Block, Ned (1995) «On a confusion about a function of consciousness», *Behavioral and Brain Sciences*, 18, pp. 227-87.
- Budd, Malcolm (1991) *Wittgenstein's Philosophy of Psychology*, London: Routledge.
- Chalmers, David (1996) *The Conscious Mind: In Search of a Fundamental Theory*, Oxford: Oxford UP.
- Descartes, René (1984) *The Philosophical Writings of Descartes: Volume II*, John Cottingham et al. (transl.) by Cottingham et al Cambridge: Cambridge UP.
- Flanagan, Owen (1992) *Consciousness Reconsidered*, Cambridge, Mass.: MIT Press.
- Güzeldere, Güven (1997) «The many faces of consciousness: A field guide» in *The Nature of Consciousness: Philosophical Debates*, Ned Block et al (eds.), Cambridge, Mass.: MIT Press, pp. 1-67.
- Kripke, Saul (1982) *Wittgenstein on Rules and Private Language: An Elementary Exposition*, Cambridge, Mass.: Harvard UP.
- Locke, John (1975) *An Essay Concerning Human Understanding*, Peter H. Nidditch (ed.), Oxford, Clarendon.
- Malcolm, Norman (1970) «Wittgenstein on the nature of mind», in *Studies in the Theory of Knowledge*, Nicholas Rescher (ed.), Oxford: Blackwell, pp. 9-29.
- McGinn, Colin (1984) *Wittgenstein on Meaning: An Interpretation and Evaluation*, Oxford: Blackwell.
- Nagel, Thomas (1974) «What is it like to be a bat?», *Philosophical Review*, 83, pp. 435-50.

²⁸ Work on this paper was funded by The Bank of Sweden Tercentenary Foundation and The Swedish Foundation for International Cooperation in Research and Education (STINT). For helpful discussions and comments, I thank Eddie Cushman, Mats Furberg, Ingvar Johansson, Sten Lindström, Peter Nilsson, Fred Stoutland, Barry Stroud and Bertil Strömberg.

- Pitcher, George (1964) *The Philosophy of Wittgenstein*, Englewood Cliffs, N.J.: Prentice-Hall.
- Searle, John (1992) *The Rediscovery of the Mind*, Cambridge, Mass: MIT Press.
- Strawson, Galen (1994) *Mental Reality*, Cambridge, Mass: MIT Press.
- Tye, Michael (2003) «Qualia», in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/sum2003/entries/qualia/>>
- Wakefield, Jerome, and Dreyfus, Hubert (1991) «Intentionality and the phenomenology of action», in *John Searle and his Critics*, Ernest Lepore and Robert van Gulick (eds.), Oxford: Blackwell, pp. 259-70.
- Wittgenstein, Ludwig (1965) *The Blue and Brown Books* [BB], New York: Harper & Row.
- Wittgenstein, Ludwig (1997) *Philosophical Investigations/Philosophische Untersuchungen* [PI], second ed., G.E.M. Anscombe (transl.), Oxford: Blackwell.
- Wittgenstein, Ludwig (1980) *Remarks on the Philosophy of Psychology, vol. 1-2* [R1-R2], G.E.M Anscombe and G.H. von Wright (eds.), Oxford: Blackwell.
- Wittgenstein, Ludwig (1981) *Zettel* [Z], second edition, G.E.M Anscombe and G.H. von Wright (eds.), Anscombe (transl.), Oxford: Blackwell.

Pär Sundström
Department of Philosophy and Linguistics
Umeå University
Sweden
<par.sundstrom@philos.umu.se>

THE MEREOLGY OF EVENTS

Robert Allen^{*}

Two cars are moving towards an intersection, one traveling east the other going north. The driver of the eastbound car runs the red light; his car and the northbound one collide at precisely noon. Call the ensuing accident High Noon. Had the driver of one of the cars braked a second earlier, their collision would have occurred later than it did (if it occurred at all). Would that slightly postdated collision, however, have been the start of High Noon?

If an event's time of occurrence is essential to it, as maintained by Lawrence Lombard, the answer is 'no'.¹ An event, according to Lombard, is a *change*, that is, the exemplifying by an object of the «dynamic» property of going from the having of one «static» property to the having of a contrary static property: it is an altering under a determinable or, in his terms, within a «quality space.»² An event occurs at the entire time that its subject is exemplifying the dynamic property involved therein: the closed interval from the last moment it possesses the about to be yielded static property to the first moment it instantiates its replacement.³ Lombard believes that this period is essential to the event, so that High Noon must commence at noon. In other words, an event such as High Noon could not have occurred without all and only the «temporal parts» it had occurring. (In «possible worlds» semantics, in any possible world in which an event occurs it has all and only the temporal parts it has in any other world in which it occurs.) Someone might simply *assume* that an accident involving the two cars in the above counterfactual situation would be High Noon rather than a distinct event. But a philosopher *arguing* for this position, he contends, would be guilty of committing a scope fallacy, confusing the true *de dicto* statement that the cars involved in High Noon could have

I am grateful to Lawrence Lombard and Cynthia Stern, who insightfully commented on the versions of this paper presented at the 1998 APA Eastern Division Meeting and the 1998 Central States Philosophical Association Meeting respectively. I also thank Linda Zagzebski, Casey Swank, Andrew Newman, Chris Swoyer, Tamoi Sagenni, Terence Parsons, and W.R. Carter for making helpful suggestions at those colloquia.

¹ This position is defended in «Sooner or Later,» *Noûs* XXIX, No. 2 (Sept. 1995) pp. 343-59. Lombard's overall view of events is presented in *Events: A Metaphysical Study* (London: Routledge and Kegan Paul, 1986) and «Ontologies of Events,» in *Contemporary Readings in the Foundations of Metaphysics*, ed. Stephen Laurence and Cynthia Macdonald (Oxford: Blackwell Publishers, 1998) pp. 289-90. Excellent discussions of his work are to be found in Greame Forbes *The Metaphysics of Modality* (Oxford: Clarendon Press, 1985) pp.207-15 and Helen Steward *The Ontology of Mind* (Oxford: Clarendon Press, 1997) pp. 58-65.

² *Events*, pp. 166-72 and «Ontologies of Events,» pp. 289-90.

³ *Events*, 1986, pp. 132-6.

collided earlier than they did with the false *de re* claim that High Noon could have begun sooner than it did.⁴

This issue may seem arcane and only of interest to event theorists. But more is at stake here than initially meets the eye. For we are wont to speculate as to what difference a temporal difference would have made: historians and journalists, e.g., often speculate as to what would have been the case had movements, battles, rescue attempts, etc. begun or ended sooner than *they* in fact did. Such speculation, if Lombard's view is correct, is metaphysically groundless. Talk of how an historical event would have «gone down» differently had *it* occurred in even a slightly different time frame could not be taken literally. It would strictly speaking have to be understood as meaning that a *sentence* whose subject designates a distinct event having the same name as it would have had a certain truth-value. Likewise, a scientist claiming that a given experiment would have yielded the same result had *it* been performed at a different time must not be taken at his word: for his experiment itself could not have been performed at any other time. Those who engage in historical speculation or scientific experimentation may be disinclined to venture an opinion on this matter, leaving its resolution to metaphysicians. But, given their habit of speaking counterfactually of events starting or ending sooner than they in fact did, it appears that they are at least implicitly committed to the denial of Lombard's view. Thus, what is at stake here is the prevalent understanding of a common practice.

I shall argue below that the key assumption in Lombard's argument against identifying an actual event with a counterfactual event having a different time of occurrence is false. Specifically, it is my intention to show that it is possible for an event to be identical with one of its proper parts. Thus, we are going to address, in the context of individuating events, the issue of whether or not identity is a relation that holds, if it holds at all, of necessity. This issue also arises, of course, in regards to the identity of material substances. Along the way, it will be necessary to speak to this matter as well as propose and defend an alternative to Lombard's criterion of event identity. In the end, we will have settled a dispute between Ockhamists and Lockians over how to characterize the relationship between a thing and its constituents.

Here is an overview of what follows. I begin by detailing Lombard's argument for temporal essentialism. I go on to point out how it relies on a questionable mereological principle, an analog of which he seems committed to denying. That is, I contrast the solution it yields to a familiar metaphysical puzzle, having to do with the relationship between a material substance and its parts, with the more plausible one entailed by its denial. I then adopt and apply the latter solution to the case of an event and its parts, showing how it supports the common practice discussed above. Next on the agenda is a discussion of Lombard's anti-atomism regarding events and how it compares with the mereological implications of my own view of the matter. The penultimate task is to formulate a suitable replacement for Lombard's criterion of event identity. I conclude by rebutting the objections that Lombard raises against my position.

We proceed to

⁴ «Sooner or Later,» p. 344.

Lombard's argument

Imagine the sinking of a ship, S. This event consists of the sinking of each one of S's (spatial) parts, these sinkings being the temporal parts of the event «the sinking of S.» Now had the last of the ship's actual parts to sink not sunk, the sinking of S would have ended sooner than it did (assuming, of course, that it had commenced at the same time as it actually did). Lombard's argument against identifying the actual and counterfactual sinkings is as follows:⁵

1. If the sinking of S that actually occurs, A, were identical to the sinking of S that counterfactually occurs, C, then A could be identical to a proper temporal part of itself, since C is A minus the sinking of the absent part (viz., the sinking of S that occurs before the last part to sink sinks).
2. It is not possible that something be identical to one of its proper parts.
3. Therefore, A is not identical to C (nor to, generalizing from 1 and 2, any event that would have ended sooner than it did- A essentially ended when it did).

If sound, Lombard's argument would show that «mereological essentialism» is true of events: «across worlds» an event could not gain or lose temporal parts just as, if this doctrine were of material substances, such an object could not gain or lose spatial parts across time. His argument is modeled on a demonstration of Peter van Inwagen that material substances are not composed of temporal parts.⁶ Let us look then at van Inwagen's reasoning, since it is going to stand or fall with Lombard's:

1. If some now defunct material substance, O, had had temporal parts, then it had a temporal part, O*, that existed as long as it did minus one minute.
2. If O had gone out of existence one minute sooner than it actually did, then it would have been identical to O*, one of its proper temporal parts.
3. But no material substance could be (counterfactually) identical to one of its proper temporal parts, something from which (in fact) it is distinct.
4. Therefore, either O had to exist as long as it did or it had no temporal parts.
5. O could have gone out of existence sooner than it did.
6. O did not have temporal parts.

Though not wishing to be counted amongst the friends of temporal parts, I think that this argument is unsound: (3), I believe, is false. More importantly, for present purposes, I believe that Lombard himself is committed to its denial. To see why, let us consider the following puzzle.⁷ How should Lombard describe the relationship between that part of a man that is left

⁵ *Ibid.*, p. 349-51.

⁶ Peter van Inwagen, «Four Dimensional Objects,» *Noûs* XXIV (1990) pp. 245-55, esp. p. 253 and «The Doctrine of Arbitrary Undetached Parts,» *Pacific Philosophical Quarterly*, vol. 62, No.2 (1981) pp.123-37 esp. pp.132-5.

⁷ Due originally to the Stoic Chrysippus. The puzzle resurfaces in David Wiggins, «On Being in the Same Place at the Same Time,» *Philosophical Review* 77: 90-105. Cf. also Michael Burke, «Dion and Theon: An Essentialist Solution to an Ancient Puzzle,» *Journal of Philosophy* 91 (1994): 123-39, Mark Hellar, *The Ontology of Physical Objects: Four Dimensional Hunks*

following the amputation of one of his limbs and the man himself (the person who underwent the operation), which now seem to be «co-located,» that is, simultaneously occupying the same space? Are they simply identical, as maintained by Ockhamists, who believe that «plurality should not be posited without necessity»?⁸ Are they distinct with the man being (only) constituted by one of his former (spatial) parts, the Lockean solution favored by those who hold that co-located objects can be distinct?⁹ Or is it a new man that emerges from the operation, the one who underwent it having perished as the result of losing one of his parts, the solution of mereological essentialists?¹⁰ Or, finally, is that part of a man that is left following the amputation of one of his limbs something that did not exist prior to that procedure being performed, it being then merely «an arbitrary undetached (spatial) part»?¹¹

In the context at hand, the last two answers may be quickly ruled out. The third is not an option for Lombard, who eschews mereological essentialism vis-à-vis material substances. The fourth is put forth by van Inwagen, who denies that things have arbitrary undetached spatial parts.¹² For Lombard, though, agreement here would be problematic, since he needs arbitrary undetached *temporal* parts to make his *reductio* work. Thus, sans an argument to the effect that events have arbitrary undetached temporal parts, even though material substance

of Matter (Cambridge: Cambridge University Press: 1990), Samuel Levy, «Coincidence and Principles of Composition,» *Analysis* 57 (1997): 1-10, Harold Noonan, «Wiggins on Identity,» *Mind* 85: 559-75, Eric Olson, «Dion's Foot,» *The Journal of Philosophy* XCIV (1997): 260-65, Michael Rea, «The Problem of Material Constitution,» *The Philosophical Review* 104 (1995): 525-52, Peter Simons, *Parts: A Study in on Ontology* (Oxford: Clarendon Press, 1987), Judith Thomson, «Parthood and Identity across Time,» *Journal of Philosophy* 80 (1983): 201-20 and «The Statue and the Clay,» *Noûs* XXXII (1998): 149-73, and Peter van Inwagen, «The Doctrine of Arbitrary Undetached Parts,» *op. cit.*

⁸ See Stephen F. Brown's «Foreword» to Ockham: *Philosophical Writings*, ed. and trans. Philotheus Boehner, Indianapolis: Hackett Publishing Company, 1990, pp. xx-xxi. Contemporary philosophers who defend this view are: André Gallois (in *Occasions Of Identity: The Metaphysics of Persistence, Change, and Sameness*, London: Oxford University Press, 1998), George Myro (in «Time and Essence,» *Midwest Studies in Philosophy* 11, 331-41), John Perry (in «Can the Self Divide?» *The Journal of Philosophy* 69: 463-880), and the present author (in «Identity And Becoming,» *The Southern Journal of Philosophy* XXXVII, 527-548).

⁹ See John Locke, *An Essay concerning Human Understanding*, ed. P.H. Nidditch, Oxford: Oxford University Press, 1975, II.xxvii.3. For a discussion of Locke's view on constitution see Michael Ayers, *Locke: Epistemology and Ontology*, London: Routledge, 1991, pp. 207-15. This position is advanced today by David Wiggins (in *Sameness and Substance*, Cambridge: Harvard University Press, 1980), Nathan Salmon (in *Reference and Essence*, Princeton NJ: Princeton University Press, 1981), Mark Johnson (in «Constitution is not Identity,» *Mind* 101, 1992: 89-105), Lynne Rudder Baker (in «Why Constitution Is Not Identity,» *Journal of Philosophy* 94, 1998: 599-621), and Judith Jarvis Thomson (in «The Statue and the Clay,» *Noûs* XXXII, 1998: 149-73).

¹⁰ The classical proponents of this solution are Peter Abelard (cf. D.P. Henry *Medieval Logic and Metaphysics*, London: Hutchinson University Library, 1972), G. W. Leibniz (in *New Essays Concerning Human Understanding*, Book II Chapter xxvii, Cambridge: Cambridge University Press, 1996) Thomas Reid (in *On the Intellectual Powers of Man*, Essay III Chapter 14, in *The Works of Thomas Reid*, ed. Sir William Hamilton, Thoemmes), and G. E. Moore (in *Philosophical Studies*, Paterson NJ: Littlefield, Adams, & Co 1959: 287-8). Amongst contemporary philosophers, Roderick Chisolm is its best known defender (in *Person and Object*, LaSalle: Open Court Publishing Company, 1976), 89-113.

¹¹ The view of van Inwagen (in «The Doctrine Of Arbitrary Undetached Parts» *op. cit.*).

¹² *Ibid.*

lack arbitrary undetached *spatial* parts, Lombard can not avail himself of van Inwagen's answer.

That forces him to choose between the first two solutions, that is, between holding that the arbitrary undetached part has become the whole of the person or that it has become merely co-located with and that which constitutes that person. Assuming that he does not wish to «multiply entities beyond necessity,» he is then left with answer one, the Ockhamist solution. And even if he lacks Ockhamistic scruples, there is a further reason why the co-location solution to this puzzle should appear unattractive to him: if we allow in the case of a material substance that the constituting and constituted are distinct objects, so that any space containing one such thing contains at least two,¹³ then consistency would seem to require us to say the same thing of an event and the events (temporal parts) «making it up». Now, while we do not think of the simple *sum* of events constituting an event as being itself an event, (lacking here the correlates of mass terms such as 'lump', 'portion', and 'piece') in the way that we do think of, say, the lump of clay making up a statue as being itself a material substance, we do take any *member* of such a sum to be an event. (On the other hand, we do not take *every* spatial part of a material substance as being a material substance *on a par with* the material substance that, in conjunction with the other members of the constituting aggregate, it constitutes. Only a «component» is treated as such.) Thus, in the case of a baseball game, we should have to say, unless we identify an event with the events constituting it, that during any one of its innings there are really two events going on, the game itself, which occurs at any time during which one of its temporal parts takes place, and the inning then being played, which we are refusing to identify with the game that it makes up at that time. For my part, though, I take it that I am seeing only one thing taking place when I am watching the sixth inning of some ballgame: the sixth inning, that is, the ballgame at that time. Anyone who shares this intuition would have to be thinking of the event constituting an event at any given time and the longer occurrence of which it is a temporal part as in some sense one.

In the case of a material substance, it is not necessary to identify any one of its proper (spatial) parts with the object itself (although, in a sense explicated below, I would identify it with the *aggregate* thereof) since it does not (simultaneously) occupy the same space as any one of its (spatial) parts: a material substance and one of its spatial parts are, to use David Armstrong's terminology, «partially identical.»¹⁴ But it seems that we *are* forced to decide whether or not to identify an event with any given one of its (temporal) parts, since, as we have seen, an event and each one of its (temporal) parts do occur simultaneously (in the same quality space). (Presently, I shall give my reasons for thinking that this matter cannot be resolved by drawing a distinction between an event's occurrence and its occurring.) To put this point another way, «I see the church» while spotting only its steeple *is* elliptical for «I see the church's steeple», whereas, «I am watching the ballgame», said during its sixth inning, is not elliptical for «I am watching the sixth inning». (We begin to see, then, that the temporal parts of an event, unlike the spatial parts of a material substance, are all one and the same, an event's relation to the times at which it occurs being like that of an *in rebus* universal to

¹³ Only in the region of space occupied by a mereological «simple» is there one and only one thing. Cf. Ned Markosian, «Simples,» *Australasian Journal of Philosophy*, vol. 76 #2 (1998): 213-28.

¹⁴ D.M. Armstrong, *A World of States of Affairs*, (Cambridge: Cambridge University Press, 1997) pp. 17-8.

its instances, an idea developed below.¹⁵) Thus, answer one, according to which «constitution is identity,» best conforms to the way in which we typically individuate material substances and the events in which they are involved. (The reason that this practice does not commit us to mereological essentialism emerges below.)

Of course, as Lombard *et al* have pointed out, there are cases in which two events occur simultaneously in the same place. That is what happens anytime a single subject changes simultaneously in more than one «quality space.»¹⁶ But cases such as a metal ball becoming warmer while it rotates are importantly different than the one discussed in the last paragraph: they do not involve events one of which is a proper part of the other, that is, occurrences that are, to again borrow David Armstrong's terminology, partially identical. Their distinctness can be made intelligible by appealing, as Lombard does, to the difference between the quality spaces involved in each event, which makes for more than one change. How, though, could we account for the distinctness of an event with one of its temporal parts when the latter just is 'how things are' with the event's subject at a given time of its progression through the quality space in which it is changing? There is only one change occurring here.

Lombard, thus, seems committed to the Ockhamist solution to our puzzle, according to which co-location entails identity. But notice what this solution implies. Unless he is willing to deny that the arbitrary undetached part in question survives the amputation,¹⁷ he is left with a case of something *becoming* identical to one of its spatial parts, refuting the key premise in both his argument and van Inwagen's. (Of course, the latter can continue to maintain that premise, since he, unlike Lombard, is not forced to accept the Ockhamist solution to our puzzle, as noted he can avail himself of the denial of the existence of arbitrary undetached spatial parts.) The time of an event, it seems, cannot be established as essential in the way that Lombard proposes. I believe that this result should not be seen as problematic, though, since there is a clear sense in which something *can* become identical to one of its proper parts. It would be our understanding of *diachronic identity* (henceforth D-identity) whereby something «perpetuates» itself by retaining its form in the process of losing an inessential part.¹⁸ Examples of such insubstantial change abound. We have the depleted

¹⁵ Thus, I commit myself to the view that an event, like a material substance of our folk ontology, endures: is wholly present at each moment of its existence. Trenton Merricks (in «On the Incompatibility of Enduring and Perduring Entities,» *Mind* vol. 104, (1995) pp. 523-31) calls this supposition «odd» but defensible: «Suppose events are property exemplifications (such as O's being red). Why couldn't O's being red be wholly present at more than one time?» Indeed. I would note further that this view avoids the dilemma Merricks poses for those whose ontology contains both enduring and perduring things.

¹⁶ In *Events*, op. cit., pp. 165-66.

¹⁷ A position advanced by Michael Burke in «Dion and Theon ...,» *op. cit.* I criticize this view in «Identity and Becoming,» pp. 4-6.

¹⁸ I discuss this notion of trans-temporal identity and what I take to be the distinct one of «identity at a time» (C-identity) and the ineliminability of both from folk ontology in «Identity and Becoming.» Per groups, heaps, lumps, and other pluralities/aggregates, it is not clear that reference thereto is a *façon de parler*, as Lombard maintains («Lombard on Allen on Lombard: Comments on Allen's Paper,» presented at the 1998 APA Eastern Division Meeting). Bertrand Russell (*The Principles of Mathematics*, London: Allen & Unwin, 1903, p. 43, 55n.) views the class as a many as a plural «object,» while Peter Simons treats it as a concrete particular (*Parts: A Study in Ontology*, Peter Simons, Oxford: Clarendon Press, 1987, pp. 144-8). Further, plural reference seems ineliminable in a proposition such as 'if there are two As there are three classes of As' (Simon's e.g.).

military unit, the rump state, the eroded dune, etc.. At this point, we need to see how this concept of identity may be integrated with the earlier claim that «constitution is identity.» That is, we must account for

Our Dualistic Understanding of Identity

Such cases involve things being identical across time despite having once been distinct. Lombard, though, is concerned with the question of when to identify events across possible worlds. There is an interesting symmetry, as it turns out, between these matters. But before that parallel is discussed, an equally intriguing dichotomy must be accounted for. Events, it is being maintained, can be identical across worlds despite being mereologically distinguishable. On the other hand, simultaneously existing material substances cannot differ in parts without being distinct: a and b are distinct if a in w is constituted at t by different parts than b in w' is at t. However, as just noted, we countenance mereologically discernible material substances being identical across time. It seems that in some cases mereological distinguishability entails distinctness and in some cases it does not.

Why this dichotomy? It stems from the fact, remarked upon by Bishop Butler, Thomas Reid, Roderick Chisolm, *et al*, that our understanding of identity is dualistic: we are not applying the concept of D-identity in all cases in which we judge that $x = y$.¹⁹ Where we are considering some object existing at a given time and some «other» (putatively distinct) object existing *at the same time*, we employ an «extensional» and temporally relativized concept of identity (C-identity) whereby @ t, $x = y$ iff they are constituted at that time by the same thing (i.e., co-located), to determine whether or not they are identical. That is why Dion and his erstwhile arbitrary undetached part are not *counted* as distinct things whereas he is to be distinguished «across worlds» from any *simultaneously* existing person constituted by something else. (This claim should not be taken to imply mereological essentialism, as it does not hold of temporally separated individuals.) By the same token, because the sixth inning of a ballgame «makes up» the game at that point we must take «them» to be one and the same in the sense of being C-identical.

So is the sixth inning the same event as the seventh? In a sense they are, since each one is the game's subject, whatever that is, changing in the quality space of «being played.» To support the claim that the sixth inning of a ballgame is the same event as the seventh, I would, thus, argue as follows:

- 1) If the sixth inning of a ballgame is not the same event as the seventh, then what occurs at the time of the sixth inning (in the quality space in which it is taking place) is not the same as what occurs at the time of the seventh (in the quality space in which it is taking place).

Reductionism, thus, seems preferable here to eliminativism. But, in any event, as Lombard notes, I can get by with arbitrary undetached parts, as in Chrysippus' puzzle. For such a case, he concedes that he «does not have (an acceptable solution)» («Lombard on Allen on Lombard: Comments on Allen's Paper»).

¹⁹ Roderick Chisolm, *Person and Object*, pp. 89-113; Joseph Butler, *First Dissertation to the Analogy of Religion* (London: 1736), reprinted in *Personal Identity* ed. John Perry (Los Angeles: University of California Press, 1975) pp. 99-105; Thomas Reid, *On the Intellectual Powers of Man*, Essay III, Chapter III, Section II. A dualistic approach to solving a philosophical problem has also recently been taken by a free will theorist. See Saul Smilansky, *Free Will and Illusion* (Oxford: Oxford University Press, 2000). Smilansky attempts to meld the truths inherent in both incompatibilism and compatibilism into a single concept of free will, maintaining that «there is no single or exhaustive notion of moral responsibility» (p.37).

- 2) But the ballgame itself occurs at both of those times (in the same quality space as that in which those parts are taking place). At each time, the ballgame's subject is changing in the quality space of (let us say) «being played.»
- 3) Thus, the sixth inning and the seventh inning are the same event (not merely parts of the same event).

So in what sense is the sixth inning of a ballgame the same event as the seventh? In the same sense as the soldiers making up a regiment before the battle are identical to the group constituting it after the battle is over: though the latter may have fewer members it is *D-identical* to the former in virtue of the nexus provided by the regiment. That is, despite any numerical difference, they are D-identical given that each one makes up the regiment *at some time or other*. As we say, it is the same regiment but not the same group of soldiers. Likewise, the sixth and seventh innings are D-identical in that each «temporarily» constitutes the ballgame of which they are temporal parts. The following transitivity principle, a modification of the classical version that accommodates our «dualistic» understanding of identity, accounts for their being D-identical:

$$(T) ((x)(y)(z) @ t, x =_c y \ \& \ @ t', y =_c z) \Rightarrow (x =_d z)$$

This principle, which is forced upon us by our desire to avoid either co-located entities or mereological essentialism, allows us to develop a logically consistent ontology without abdicating either one of the above concepts of identity.

Mereologically, then, an event must be treated like a universal *in rebus*, being related to the distinct times at which it is occurring in the same way as the latter is related to the disjoint spaces at which it is located. It follows that we should not expect its identical instances either to begin simultaneously or to obey Leibniz's Law (unless it is temporally relativized). Compare: «The blueness of my tie is to the right of and occupies a narrower space than the blueness of my shirt pocket» vs. «The sixth inning started earlier and lasted longer than the seventh». And, just as there is no limit to the number of instances a universal *in rebus* can have, it is possible for an event to increase/decrease its duration by adding to/subtracting from its actual number of occurrences. (Here I am reminded of the relationship between a TV series and its episodes.) Thus, pace Lombard, who contends that an event *as a whole* does not occur but, rather, is only occurring when any one of its temporal parts occurs,²⁰ an event itself occurs *and* is occurring at any time at which its subject is exemplifying the dynamic property involved therein, just as a universal *in rebus* is instanced at any and all of its spatial locations.

Let us digress here to consider the mereological implications of drawing a distinction between the time of an event's occurrence- its duration- and the times of its occurring- the duration of any one of its proper temporal parts. In particular, we must determine whether or not it allows Lombard to avoid having co-located entities in his ontology. Even if it were true that a *composite* event is only occurring at any time at which one of its proper parts occurs, so that they are just partially identical, what are we to say of an atomic temporal part of an event, one whose duration is no longer than what is temporally required for a change in the relevant quality space to occur and, thus, has no proper part at which it is only occurring? Here Lombard's occurs/is occurring distinction would be inapplicable. Thus, unless one assumes an anti-atomistic view according to which an event must be composed of proper parts

²⁰ «Lombard on Allen on Lombard: Comments on Allen's Paper,» and *Events*, pp. 132-6.

that are the events that make up its occurring, which entails that the occurring of an event is always a proper part of its own occurrence when any event of which it is a part would be occurring, we would have two events *occurring* simultaneously in the same quality space were we not to identify an atomic temporal part of an event with the event of which it is a part, since the former occurs as it is occurring. This co-location of occurrences would be as ontologically profligate as the co-location of occurrences Lombard’s distinction would obviate.

Lombard avoids having to answer this question by advancing the just mentioned anti-atomistic view of events. He is forced to accept this position, independently of his concern to avoid co-located occurrences, by his handling of the problem of direct change, that is, going from having one property to having of another property without having an intermediate property. Such a change would take place in a «discrete» quality space: one in which there are pairs of properties between which there lies no intermediate properties. The quality spaces in which one would go from having black to gray hair and one to two strikes are examples of this type of quality space. The following argument, however, seems to show that it would be impossible for changes to take place therein:

(1)	Time is dense; between any two instants there are other times	(Assumption)
(2)	If object <i>o</i> goes from having <i>P</i> for the last time at <i>t</i> to having <i>Q</i> for the first time at <i>t'</i> , then there is a time <i>t*</i> between <i>t</i> and <i>t'</i>	(1)
(3)	If <i>o</i> 's change from <i>P</i> to <i>Q</i> could be direct, then there is no property within the <i>P/Q</i> quality space that <i>o</i> has at <i>t*</i>	(Definition of direct change)
(4)	<i>O</i> goes from having <i>P</i> for the last time at <i>t</i> to having <i>Q</i> for the first time at <i>t'</i>	(Assumption)
(5)	There is a time <i>t*</i> between <i>t</i> and <i>t'</i>	(2, 4)
(6)	If there is a time <i>t*</i> between <i>t</i> and <i>t'</i> , then there is a property within the <i>P/Q</i> quality space had by <i>o</i> at <i>t*</i>	(Assumption)
(7)	There is a property within the <i>P/Q</i> quality space had by <i>o</i> at <i>t*</i>	(5, 6)
(8)	<i>O</i> 's change from <i>P</i> to <i>Q</i> could not be direct	(7)

Lombard responds to this argument by noting that if a direct change could occur at an instant, instead of requiring more than one instant to occur (in effect, if 4 were false), it would not present a problem, since then there would be no need to account for *o*'s status within the *P/Q* quality space during the interval between *t* and *t'* that must exist if time is dense.²¹ Given, however, that a change is a «process,» a «transition» from one property to another, there arises the problem of explaining the possibility of direct change. Lombard’s solution is to maintain that «such (a change consists) of other events that are dense changes in (a dense

²¹ *Events*, p. 141.

quality space).» That is because all the «ultimate» quality spaces in which events occur are dense, with no direct changes taking place therein.²² Absolute atomic events are thus ruled out. Instead, atomic events are, relative to a scientific theory T, (and leaving out several details not germane to the present discussion) those temporally continuous events involving changes to T's atomic objects occurring within T's ultimate quality spaces.²³

There are several problems with this approach. To begin with, it saddles Lombard with anti-atomism regarding events. He believes that, «short of trying to take (seriously) the idea of an instantaneous event,» maintaining this view is unavoidable.²⁴ I presently shall show that this dilemma is false. The second problem with it is that Lombard offers, in his own words, «no direct, independent argument» for the claim that all events are composed of dense changes.²⁵ It is, thus, merely an expedient for solving the problem of direct change. Sans such an argument, then, it should be abandoned if a more plausible solution to that problem can be found. Thirdly, since an argument parallel to the one just given can be constructed to show that physical objects cannot touch, Lombard must accept that conclusion as well. Though he is willing to do so, we have here another move that had best be avoided if possible.²⁶ In the case of events, he does not think that it can be eschewed without allowing that events can occur instantaneously. But that cannot be right, since there are other premises in the above argument one can challenge.

Finally, and most importantly, Lombard does not so much account for direct change as deny its occurrence. Direct change only appears to occur, on his view, since the difference between the quality space in which it would occur and the quality space in which an indirect change would occur precludes him from *identifying* it with an indirect change (given his criterion of event identity discussed below). What *really* happens in such a case is a series of indirect changes. His position here is eliminativist rather than reductionist: it is akin to the view that the posits of «folk ontology» are nothing more than collections of sub-atomic particles; they do not exist, so that the question of whether or not they are identical to such aggregates need not arise. When Lombard says that direct changes are «really composed» of

²² *Ibid.*, p. 142.

²³ *Ibid.*, p. 168-7.

²⁴ *Ibid.*, p. 261, note 18.

²⁵ *Ibid.*, p. 140.

²⁶ *Ibid.*, p. 138. Here is that argument:

1. If o and o' touch, then there must be a place, p, at which they come in contact
2. Two objects cannot simultaneously occupy the same place.
3. Thus, there is no place p occupied by both o and o' (1,2)
4. o occupies p
5. Thus, o' occupies p' and $p \neq p'$ (3,4)
6. Space is dense; between any two spaces there are other spaces
7. Thus, there are spaces p'', p''' between p and p' (5,6)
8. Those spaces are occupied by neither o nor o'.
9. If there are spaces between the spaces occupied by two objects, they do not come in contact
10. Thus, p and p' do not come in contact. (8,9)
11. Thus, p and p' do not touch (e.g. my hand can not touch my computer's keyboard) (1,10)

indirect changes, he cannot mean (something analogous to) what the reductionist regarding material substances holds, for the latter accepts the reality of ordinary material substances, she simply believes that they are identical to that of which they are constituted. I should think that a solution to the problem of direct change that takes seriously the possibility of its occurrence would be preferable to one that treats it as an illusion.

I can make out two alternatives. First, there is the option of taking the above argument as a *reductio* on the thesis that time is dense, which, after all, is arguably not a part of common sense, unlike the belief in direct change. (And even if time were intuitively considered dense, this response would be more plausible than Lombard's, which makes direct change only apparent *and* requires atomic events to have parts, since its advocate must abdicate fewer intuitions than Lombard is required to give up.) If time were discrete, then there would be no need to account for *o*'s status in regards to the P/Q quality space during the interval between *t* and *t'*. We could hold that, since the change is direct, the only times involved are *t* and *t'*: they would be the times at which the event began and ended respectively as well as the times during which it took place, making it, in Lombard's terms, both an occurrence and an occurring. It would, thus, be an event having no *proper* parts at which it is occurring, raising the question Lombard had hoped to avoid of whether or not its occurring would be identical to the simultaneous occurring in the same quality space of any event of which it were an atomic part.

A more plausible option still would be to deny premise 6. Why should an object changing in a quality space exemplify a quality belonging to that space at *every* instant at which that event is occurring? If the quality space in question is dense, then, assuming time is also dense, there *would* be a one-to-one correlation between the instants that make up the interval at which the event occurs and properties that make up the quality space. But such a correlation would not obtain between the instants that make up a dense period of time and the properties composing a discrete quality space. Thus, if *o* goes from P at *t* to Q at *t'* in a discrete quality space, the instants between *t* and *t'* (which we are supposing to be infinite) must be times at which *o* is devoid of any properties belonging to the P/Q quality space. To suppose, as Lombard does, that between *t* and *t'* «(*o*) must have a quality in (such a) space»²⁷ seems tantamount to denying the existence of discrete quality spaces, question begging in the present context. Real, not merely illusory, direct change is, thus, possible. In the bargain we secure atomic events, since a change from P to (its neighbor) Q in a discrete quality space will not break down into the occurrence of further events making up its occurring. Since it does not have a proper part at which it is occurring, it will occur at the same time as it is occurring. Thus, the question will again arise as to whether or not as an occurring it is identical to the simultaneous occurring of any event of which it is an atomic part. In the context of individuating events, therefore, the occurs/is occurring distinction cannot help Lombard avoid the dilemma of accepting either co-location or the identity of a thing with one of its parts.

Picking up where we left off before our digression, we must now formulate

An alternative to Lombard's Criterion of Event Identity

To individuate events across worlds we cannot rely on the notion of C-identity, since we are always dealing here with things taking place at different times: as in the case above in which the actual collision occurs at noon and the counterfactual one at noon plus one second.

²⁷ *Ibid.*, p. 144.

(Moreover, since an event, as we have just seen, does not have full-fledged parts (each one existing independently of the whole to which it belongs) it does not even make sense to ask of one, as it would in the case of a material substance, whose parts could exist in the absence of that which they make up, «could it be identical to an event constituted at some time by an event that is distinct from the event constituting it at that time?» — that would be a different event, since an event (in effect) constitutes itself at any time at which it occurs.) Thus, we are in need of a criterion of cross-world event identity analogous to the criterion of diachronic identity for material substances sketched above.²⁸ I propose the following (relying on Lombard's conception of an event as the changing of an object in a quality space):

(EI) Necessarily, E in w is identical to E' in w' iff a) the subject of E is identical to the subject of E' (or at least is a «counterpart» of), b) E and E' are changing in the same quality space and c) (parting company with Lombard) the time at which E occurs includes the time at which E' occurs or vice-versa

EI gives us the desired result that an event could occur during an interval other than its actual time of occurrence. As stated, however, it is too narrow, since if an event could have begun earlier/later than it did and an event could have ended sooner/later than it did, then an event could have begun *and* ended earlier/later than it did, an event could have begun earlier than it did *and* ended later than it did, and an event could have begun later than it did *and* ended sooner than it did. To accommodate these intuitions, we must modify clause c of EI to read «the time at which E occurs overlaps the time at which E' occurs».

The following case appears to show that EI is also too liberal. In w, object o covers the distance between point A and point B during the interval from t^1 to t^6 ; in w', o journeys from A to B between t^1 and t^2 , remains at B until t^3 , swiftly returns to A by t^4 , where it stays put until its departure at t^5 for B, which it reaches at t^6 . By EI, the distinct passages of o in w' are identical to its single journey in w. This supposed defect could be removed by adding a fourth clause to EI:

d) there is not an event E'' that also meets a, b, and c in regards to one of either E or E' while not being identical to the other

This move, though, in addition to appearing *ad hoc*, entails the denial of David Wiggins' «only a and b rule,» stating that whether or not $a = b$ is independent of what else exists. Since I am not comfortable with this position in regards to material substances, I am not inclined to adopt it here. Moreover, since I have already defended the contingency of identity in the case of material substances, consistency would seem to require me to treat the above case as an instance of (the possibility of) one *event* becoming two. Consider these others: in w, the Hundred Years War is interrupted by a one hour armistice; in w, I fly to Paris twice as fast as I actually did, realize upon touchdown that I left my luggage at the departure gate, immediately hop on a flight back to Detroit (travelling at the same high speed), retrieve the bags, and return as swiftly as I came. I would prefer to treat these cases as analogous to cases of «fissioning» material substances (as with an ameba dividing or so called «brain bifurcation»), the least implausible response to which being, as noted above, to deny the necessity of identity.

²⁸ Concerning physical objects we should have to say that necessarily, any two of them existing non-simultaneously in separate worlds are identical iff one is how the «other» would have perpetuated itself had it existed therein, although specifying the conditions under which x would have perpetuated itself as y would be tricky.

Sans a principle of unity for events ruling it out, EI implies, e.g., that the French Revolution could be identical to the Battle of Dienbeinphu. If the former had occurred for several month longer, it would have overlapped temporally with the Terrors, making them, according to EI and the thesis that the temporal «parts» of an event are identical, the same event (assuming that they share a subject and take place within the same quality space). But, then, by successive similar extensions of this event, an event that overlapped temporally with the Battle of Dienbienphu could be formed, making it and the French Revolution «temporal parts» of the same event and, thus, the same event (again assuming sameness of subject and quality space). T, propounded above, allows us to account for this result, just as it establishes that the innings of a ballgame are D-identical. (Indeed, a historian studying 19th and 20th century Europe would be inclined to posit the existence of just such an event: French history from the Revolution to Dienbienphu.)

I conclude by responding to

Two objections that Lombard would raise against my rebuttal

First, it is open to him to reject C-identity, to deny that constitution is identity. In fact, that is precisely his position. Before examining the arguments he puts forth in its favor, it is worthwhile asking whether or not it is a view with which someone with Ockhamistic scruples could reconcile herself. Does it leave her with any way of avoiding the awkward conclusion that our world contains vastly more entities (material substances and events) than ordinary inventorying would indicate? One might try to distinguish here between the notions of being an object and being a (countable) individual, refusing to grant constituting entities the latter status.²⁹ But there does not seem to be any «conceptual space» between these ideas: it is not as if we never count portions of clay, hunks of wax, n-numbered collections of molecules etc.. Thus, those who deny that constitution is identity seem to be committed to biting the above bullet.

In the case of material substances, Lombard argues that the constituted and the constituting are not identical because: i) there can be a time when the one exists and the other does not, ii) at that time they would not be identical (since iii) a and b cannot be identical at a time at which one exists and the other does not) making them non-identical, since iv) if there is any time at which a and b are distinct then there is no other time when they are identical.³⁰ But a defender of the «dualistic» view of identity sketched above would take each one of the previously noted cases of mereological alteration as a counterexample to (iv). The depleted military unit, rump state, and eroded dune is each C-identical to something from which it was once distinct by that same criterion. That is because the sums of parts to which each persistent has been «occasionally» identical are relata of another equally indispensable identity relation: D-identity. Lombard notes up front that his argument here «does not address the concerns of believers in relative (occasional) identity» and indeed it does not.

Lombard's argument against identifying an event with the temporal parts constituting it relies on modal considerations.³¹ No temporal part of an event is essential to its occurrence:

²⁹ Arda Denkel defends this view in *Object and Property*, (Cambridge: Cambridge University Press, 1997) pp. 89-90.

³⁰ *Events*, op. cit., pp. 250-52.

³¹ Cf. «Events and their Subjects,» *Pacific Philosophical Quarterly* Vol. 62 (1981): pp. 138-47.

had its subject had different spatial parts it could still have taken place, though it then would have had different temporal parts (owing to its subject having different spatial parts than the ones that are themselves the essential subjects of the events that are its temporal parts). On the other hand, the sum of an event's temporal parts, which is what constitutes it, could *not* occur unless each one of those parts did. This difference between their persistence conditions makes for the distinctness of an event and the sum of temporal parts of which it is constituted. Here I would point out, though, that the concept of an entity's essence should not come into play in the making of judgments of synchronic identity. Since it is supposed to tell us how an entity would behave across time- what sorts of changes it could and could not endure- its proper role is in facilitating the determination of whether or not there is «continuity under a sortal» between temporally separate entities. It is, thus, a category mistake to apply this notion, as Lombard does, in a context in which one does not need to know what an entity would have been like in the past had it *then* been identical to something, that is, where one's identity concerns extend only to entities existing simultaneously, such as an event and the sum of its temporal parts.

At this point, Lombard would no doubt question my claim that our understanding of identity is dualistic. It cannot be denied, however, that common sense works with more than one notion here: witness our embrace of both Okhamistic counting and continuants capable of mereological alteration. We would have to abdicate the former were we to repudiate C-identity; the latter would be a casualty of giving up D-identity in favor of mereological essentialism. We have shown, moreover, how these concepts can cohere. To avoid contradictions, all that is required is a refinement of the principle of the transitivity of identity, as formulated above. Letting go of the idea that an event's time is of its essence, on the other hand, costs common sense nothing, since, as Lombard himself concedes, it is not a part of that body of beliefs to begin with.

Conclusion

We began by establishing that those who deny the possibility of a thing becoming identical to one of its parts and accept the possibility of mereological alteration are faced with the dilemma of either denying the existence of arbitrary undetached parts or accepting co-located objects and events. Preferring the sparsest possible ontology, we embraced a dualistic conception of identity that allows for trans-temporal/modal numerical sameness of objects/events despite the changing of their parts while ruling out the sharing of a spatial/temporal location by more than one object/event. Mereologically distinct entities standing in the relation of D-identity, that is, obeying our transitivity principle T, form a continuant that is C-identical with its constituent(s) at any given time. Thus, we accomplished our main objective: refuting Lombard's essentialism regarding the time of an event, thus allowing for the adjustment of an event's temporal parameters in historical and scientific speculation. It is possible for an event to have taken place in a period of time that is longer/shorter than the one in which it actually occurred; i.e., to have had more/fewer temporal parts than it had in actuality. Out of this critique grew a novel view of the relation of an event to the times at which it occurs. An event turns out to be akin to an *in rebus* universal, enduring across time as the latter distributes itself throughout space. Admittedly, this view entails the counterintuitive result that temporally separated events such as the sixth and seventh innings of a baseball game are in some sense one. This single drawback must, however, be balanced against the conceptually troublesome moves of my opponent: his revision of the semantics of counterfactual discourse and embrace of co-locationism. Assuming

that it is not possible to solve a philosophical problem without modifying one's system of beliefs, one should strive instead to preserve as many intuitions as possible. Moreover, the notion that temporally separated events are identical appears less strange when one considers our tendency to unify the stages of a material substance's career. Temporal parts theorists or «four-dimensionalists» suggest treating material substances as events.³² Here the opposite tack is taken: a diachronic principle of unity for material substances is applied to events. It should also be kept in mind that this principle co-exists with a concept that allows for judgments of distinctness in cases of mereological alteration, their applications being a function of temporal focus. Following this program, to be sure, requires cognitive nimbleness, given its bifurcation of the meaning of identity. Yet, given our willingness to accept shifting views of reality, it may be promoted as user-friendly.

Robert Allen
9901 Marion
Redford, MI 48239-2015
<allen1rf@mail.cmich.edu>

³² See Heller, *op. cit.* and Theodore Sider, *Four-Dimensionalism* (Oxford: Oxford University Press, 2001) for expositions of this view.

SORITES (ΣΩΠΙΤΗΣ), ISSN 1135-1349
http://www.sorites.org
Issue #16 — December 2005. Pp. 38-45
Dismantling the Straw Man: An Analysis of the
Arguments of Hume and Berkeley Against Locke's
Doctrine of Abstract Ideas
Copyright © by SORITES and Rhys McKinnon

DISMANTLING THE STRAW MAN: AN ANALYSIS OF THE ARGUMENTS OF HUME AND BERKELEY AGAINST LOCKE'S DOCTRINE OF ABSTRACT IDEAS

Rhys McKinnon

Both Berkeley and Hume made an effort to respond to Locke's theory of abstract ideas: Berkeley vehemently rejected it, and Hume offered arguments in further support of Berkeley. Berkeley was quick to label Locke's product as an «abuse of language.»¹ Hume was not nearly as severe in the language that he used in his arguments in support of Berkeley against Locke, but believed that he was entirely in agreement with Berkeley's own arguments. However, there are grounds on which to question the strength of both Berkeley and Hume's attacks on Locke's account of abstract ideas: both Berkeley and Hume, possibly unbeknownst to them, reach some conclusions that *agree* with Locke. A journey through a close exegesis of Locke, and the subsequent attacks made by Berkeley and Hume, will lead to a conclusion that many of the attacks were merely of a straw man² nature and that many of their conclusions (Berkeley and Hume) can be taken to be in agreement with Locke, instead of their believed stance of repudiating Locke. This paper is not a comparative analysis of the three philosophers' conceptions of abstract ideas; instead, it is an analysis of the attacks of Berkeley and Hume on Locke, given a close reading of Locke (as well as, obviously, Berkeley and Hume).

Locke is quick to state that humans have a need for abstract (general) ideas and terms: they allow for easier communication of ideas³ Hume would certainly agree that the «custom» of abstract general ideas and terms facilitates communication⁴ but Berkeley certainly would not.⁵ Locke argues this from the position that it is not very useful to attempt to communicate concepts with another using only particular cases, since it requires the other to be «acquainted

¹ Berkeley, *Principles of Human Knowledge*, Introduction, 6; p38.

² «Straw man» referring to the logical fallacy of attacking a position that someone never actually offered.

³ Locke, *Essay Concerning Human Understanding*, 3.3.3.

⁴ *Treatise*, 1.1.7.7-8, for example.

⁵ *Principles*, Introduction, 14.

with all those very particular Things» which are being used for reference⁶ Such a statement is what Aristotle concluded in his *Metaphysics*: experience is of particulars, but *knowledge* is of universals⁷ Although identifying the utility for abstract ideas is necessary⁸ (so that it is established as to *why* one should be concerned with the topic of the meaning and genealogy of abstract ideas), Locke's main focus is on the genealogy of abstract ideas/terms and what their specific meanings are. Locke makes a distinction between *ideas* and *terms/words*. Ideas are the objects of the mind when we think⁹ and terms/words are the sounds/written symbols that we use to signify (use as signs) certain ideas in our mind.¹⁰ Thus, ideas must be ontologically prior to terms, if the term is to have meaning *qua* being a sign for an idea.

Locke, as well as Hume¹¹ believes that ideas must come from either *sensation* (experience of the senses) or *reflection* (reasoning/«internal sense» of the mind.¹² In the genealogy of abstract ideas, and their general terms that signify the ideas, do the ideas come from sensation or reflection? Locke argues that there only exist *particular* things in the world, not general (abstract) things; thus, there cannot exist an object in the world from which the abstract idea can be arrived at strictly through the senses.¹³ Therefore, if there are any abstract ideas, they must be from reflection. Ideas are made general (abstract) by separating from the particular idea all of the determinate qualities (*viz.* time, place, colour, size, etc.) that differentiate that specific idea from others that are similar to it (which one would group under the same abstract idea.¹⁴ However, one cannot abstract *all* of the idea's qualities out of the particular instance: one must retain the defining qualities that make that idea of the sort that that idea belongs. An example will illustrate this last point. If one is to abstract from the idea of a specific triangle (a right triangle with vertical length of 3cm, horizontal length of 4cm, and hypotenuse length of 5cm), one can separate the ideas of determinate length and colour, but not of it being a «plane figure with three sides.» One cannot remove the last qualities

⁶ *Essay*, 3.3.3.

⁷ Aristotle, *Metaphysics*, 1.2, 982a.

⁸ This is, certainly, not a conclusive argument, as Berkeley argues (see note13); but it may have weight if the particular (thing) that the other person is acquainted with is not sufficiently similar to that of the person who is offering their own particular (thing) in their discussion.

⁹ *Essay*, 1.1.8.

¹⁰ *Ibid.* 3.1.2.

¹¹ *Treatise*, 1.2.1, though Hume makes an even more specific distinction between «ideas» and «impressions» over and above Locke's «loose» (in Hume's mind) use of the word «idea». Thus, in one sense, Hume would be discussing impressions, and another ideas; while Locke is only referring to ideas (as they have different definitions of what are ideas, and how they are made).

¹² *Essay*, 2.1.2-4. Although, in the *strictest possible sense*, Locke and Hume differ in their accounts on what is meant by Sensation and Reflection, and how ideas are made/received, I believe that their respective meanings are certainly close enough to allow the equivocation of their two sentiments about this issue.

¹³ *Ibid.* 3.3.6, but more forcefully and explicitly in 3.3.11.

¹⁴ *Ibid.* 3.3.6.

since they are what make a triangle a triangle: they are its essential qualities. Locke defines abstract ideas in terms of these essences: every distinct abstract idea has a distinct essence;¹⁵ but, he makes a distinction between supposed *real* essences, and *nominal* essences. His *nominal* essences are those in which we define the essence of a thing by a particular *grouping* of qualities which are defined by ourselves: such essences are nothing but «the uncertain and various Collection of simple *Ideas*» which our minds put together.¹⁶ *Real* essences, conversely, are epistemologically inaccessible for us, and are the characteristics of its *substance* that makes it that particular thing and not another.¹⁷ this is known as substance theory.¹⁸ Interestingly, Locke's conception of the nominal essence is what Berkeley and Hume would call the *essence*¹⁹ of a thing: a bundle of perceptions that are constantly grouped together in our minds; this is known as bundle theory. In fact, Locke thinks that when we create general ideas, and terms for those ideas, we cannot be using the real essences of a thing. If we were to attempt to use the real essences, of which we have no knowledge, then we would not be able to precisely determine when a thing ceases to be a horse, and begins to be lead;²⁰ thus, to posit the existence of real, epistemologically inaccessible, essences «is so *wholly* useless» in our conception of abstract ideas.²¹ However, if we define the abstract idea of a thing, *viz.* the abstract idea of *horse*, by using nominal essences, we are merely defining the thing (the abstract idea) by a group of simple ideas put together by the mind; in fact, Locke equates *nominal essence* («general natures») with abstract idea.²²

The main point of contention between Berkeley and Hume, on the one hand, and Locke, on the other, is the separability of ideas and qualities. In Hume's epistemology, he makes the bold claim «that there are not any two impressions which are perfectly inseparable.»²³ Thus, he claims that all things that are different are distinguishable, and all things that are distinguishable are separable by the mind (through the use of reason.²⁴ However, this is not

¹⁵ *Ibid.* 3.3.14, 28.

¹⁶ *Ibid.* 3.3.14, 25-27.

¹⁷ Woolhouse, R.S., *The Empiricists*, 85-6.

¹⁸ Dicker, G., *Hume's Epistemology and Metaphysics*, 15.

¹⁹ Hume discusses this in terms of substance in 1.1.6. His definition of substance is first merely what Locke would call the nominal essence of a thing, but that this collection of qualities are somehow inseparably connected (by either contiguity or causation). Thus, I do not find it incorrect to use the word «essence» in the sense that I have in this sentence. Berkeley, however, never spoke in terms of either essence or material substance; thus, his inclusion in this sentence was only because of the way that he defined what it is to be a certain thing: to be a collection of a certain set of ideas.

²⁰ *Essay*, 3.3.13.

²¹ *Essay*, 3.3.17, his emphasis.

²² *Ibid.* 3.3.9, lines 7-10.

²³ *Treatise*, 1.1.3.4.

²⁴ *Ibid.* 1.1.7.3.

entirely true: both Berkeley and Hume wonder if there are certain things (ideas or qualities) that are, *strictly speaking*, inseparable. The body of Berkeley's attack on Locke's doctrine of abstract ideas (and, consequently, Hume's furthering of Berkeley's arguments), is his argument for the inseparability of certain qualities. Since all three (Locke, Berkeley, and Hume) are arguing for bundle theory²⁵ the pith of a given abstract idea is a collection of *simple ideas*: it itself is a complex idea (*qua* an amalgamation). However, it is in how Locke claims abstract ideas are made that the dispute lies: can we generate an idea of a thing without a definite value to each quality, *viz.* colour, size, motion? Berkeley and Hume, *strictly speaking* (loosely speaking, we will see that Hume offers another answer), do not believe that this is possible. Berkeley argues that there are certainly some parts or qualities that one can abstract from a given thing *if and only if* those parts or qualities can exist by themselves.²⁶ For example, it is possible for one to abstract the idea of *nose* from a man since it is readily conceivable that a nose can exist without the rest of the body.²⁷ Therefore, Berkeley argues that the idea of a *Centaur* (an example akin to Locke's unicorn and mermaid examples; *Essay*, 3.3.19) is intelligible to the mind, even though one may never have *experienced* one (as the mind alone is capable of putting the two ideas together in a non-contradictory way.²⁸ However, Berkeley and, subsequently, Hume believe that there *are* certain qualities that cannot exist apart from others; these include extension, colour, and motion. Berkeley argues that colour cannot exist on its own: it cannot exist without figure, which cannot exist without extension and motion.²⁹ However, Hume believes that ideas *can* be abstracted in the mind such that we can distinguish between different *aspects* of «inseparable» qualities.³⁰ To do this is to make a «distinction of reason».³¹

A distinction of reason can be made when one considers two objects that are similar in one quality, but differ in (often all) others. Thus, the process is to discover different relations and resemblances of different sorts. Hume's example is the comparison of the following objects: a globe of black marble, a globe of white marble, and a cube of white marble.³² If

²⁵ Only when discussing abstract ideas. Locke argues for substance theory with respect to material objects, where Berkeley argues for bundle theory. Hume, however, neither commits himself to bundle theory nor substance theory: he is skeptical about substance theory, but does not want to dismiss its truth. Thus, he argues for bundle theory as the only epistemologically tenable theory, given our current state of epistemology, but not *explicitly* for the «truth» of bundle theory.

²⁶ *Principles*, Introduction, 10.

²⁷ I say this keeping in mind that some philosophers, Aristotle comes to mind, would argue that this conception is not of a nose *qua* nose because the idea of a nose is only when it is functioning as a nose (therefore it must still be attached to the body). However, it is neither inconceivable nor contradictory to imagine a nose removed from a man, even if we should call it a name other than «nose.»

²⁸ *Ibid.* 10.

²⁹ *Ibid.* 7.

³⁰ *Treatise*, 1.1.7.18.

³¹ *Ibid.* 1.1.7.17-18.

³² *Ibid.* 1.1.7.18. I believe that this example is best taken if we consider these three objects as our only experiences of objects.

one were to merely observe and consider the white globe of marble, one would not be able to separate the colour from the figure; in fact, one would not even have a concept of the abstract idea of *colour* itself since one would only be acquainted with «a white colour dispos'd in a certain form.³³ However, if one were to consider the white globe in relation to the black globe, the differing ideas of colour have a resemblance (to each other) that can lead to the creation of an abstract idea of colour through this distinction of reason. Furthermore, if one were to consider all three objects together, one would notice that there are now two separate resemblances, *viz.* colour and figure/form, in what, at first, seemed to be single and inseparable.³⁴ To speak of the continuing inseparability, though we are able to reason an *apparent* separability through this distinction of reason, when we consider the white globe and its colour and figure, to do so requires a *tacit* reference to the black globe and white cube (or other objects that allow the distinction in reason through the two different sorts of resemblance, *viz.* that the objects only differ in one aspect from the original white globe: one with the same shape, but different colour; the other with the same colour, but different shape). Hume is here proffering the belief that simple ideas, though inseparable in *reality*, can have, in a sense, differing degrees of resemblance. This resemblance allows the two simple ideas of black and white to be grouped together and seen as – only to a degree, since colour and figure are really inseparable – closer in resemblance than between either (white and black) and roundness.³⁵ Finally, Hume believes that such a process gets better with practice³⁶ and is, therefore, in agreement with the similar statement that Locke makes.³⁷ Berkeley, however, is in stark disagreement that the ability to abstract is developed since he does not *seem* to show a belief in the ability to make abstract ideas *qua* distinction of reason.

Thus, we have seen that Hume, while arguing in *support* of Berkeley, goes beyond Berkeley and, as a result, seems to be clearly agreeing with Locke in an important sense. Hume takes great care to show that although we can make a distinction of reason, there is no *real* separation of such qualities (one cannot *truly* separate colour from extension). This distinction is a point of departure between Locke and Hume: although Locke *also* argues for the *real* inseparability of ideas *in the objects themselves*, he still believes that qualities such as colour and figure can be separated from each other in the *mind*; however, Hume does not believe in the ability to really separate qualities *in the mind*. Thus, Hume, *strictly speaking*, does not believe that the qualities of colour and extension are actually being separated through a distinction of reason; it is only an *illusion*: to think of colour without extension is impossible.³⁸ However, we can see that there are different «sorts» of resemblances and, of these different sorts, colour and figure are members.

³³ *Loc. cit.* Although it is interesting that he is using language that is meaningless if we only have the experience of this one object.

³⁴ *Loc. cit.* It is important to note that although these qualities now appear separable, they still, in reality, are not (according to Hume).

³⁵ This is an expansion of part of the Appendix to volume 3 but was inserted into Book 1 found on page 18 of Norton, 2002.

³⁶ *Treatise*, 1.1.7.18.

³⁷ *Essay*, 4.7.9.

³⁸ *Treatise*, 1.1.7.18.

Finally, the other major point of contention between Berkeley and Hume, on the one hand, and Locke, on the other, is the nature of the abstract ideas themselves. As established above, Hume argues that we can have abstract ideas *qua* distinction of reason, but not abstract ideas *qua* qualities being *actually* separable (in the mind). In Locke's account of the genealogy of abstract ideas, as discussed above, abstract ideas are made by removing³⁹ all of the determinate qualities that differentiate a given particular thing from another, but leaving the qualities that the two particular things have in common as defined by the *nominal essence* for the *abstract idea*, of which these two things are members (all done in the *mind*, not in the things themselves). Thus, in the abstract idea of «human», the determinate qualities of size and colour, among others, are removed. Berkeley and Hume attack this contention. Berkeley argues as follows.⁴⁰ The abstract idea of *man* (human) must include all particular cases of humans. Since humans are of different colours, the abstract idea cannot have a determinate colour; the same applies to height, shape, and other such qualities. However, Berkeley cannot conceive of what such an idea would be since he cannot imagine an uncoloured person (more specifically, a person with indeterminate colour). Hume expands on this argument as follows.⁴¹ Hume's argument is, in my opinion, a psychological account of what it is to consider an idea. Firstly, it is impossible for one's mind to conceive of any quality without having a precise degree. When one considers a line, it has a certain length; when one thinks of a triangle, it will have *determinate* lengths and angles thus defining its shape. Furthermore, when one thinks of a triangle, it must have *both* determinate shape and lengths; one cannot exist separated from the other (this is part of Hume's argument for the *real* inseparability of qualities even in the mind). Thus, this inability to conceive of a triangle without determinate qualities implies a contradiction to Locke's abstract idea of a triangle without any determinate qualities (other than the defining qualities of being a plane figure with three sides). *Secondly, Hume argues that it is impossible for a triangle to exist which «has no precise proportion of sides and angles.»*⁴² *Furthermore, it is impossible for an idea that has both quality and quantity but no precise degree in either.*⁴³ *Therefore, according to Hume's epistemology, viz. if one cannot have an impression of a thing then one cannot have an idea of it, we cannot have any abstract ideas qua Locke's conception of abstract ideas.*⁴⁴ Interestingly, these arguments come from a fundamental misunderstanding of Locke; thus, a critique of these arguments will show that they are attacking points that Locke never made.

³⁹ I have here, and hereafter, used «remove» (and its cognates) to describe Locke's genealogy of abstract ideas. This is a result of a semantical interpretation of *Essay*, 3.3.8. He speaks in terms of: «retaining only those Qualities», «uniting them», «leaving out», and not creating new abstract ideas by «addition.» Thus, I take «leaving out the shape, and some other Properties» and «retaining» a specific set of others to mean «removing» the qualities not wished to be «retained» in the abstract idea. The use of «remove» seems to be consistent with the overall meaning of this section (3.3.8).

⁴⁰ *Principles, Introduction*, 9-10.

⁴¹ *Treatise*, 1.1.7.2-6. The most pertinent sections for our current discussion is 2, 3, and 6.

⁴² *Ibid.* 1.1.7.6.

⁴³ *Loc. cit.*

⁴⁴ It is of paramount importance to remember that this is only Hume's argument, *not* a critique of it.

A close reading of Locke finds him explicitly stating that he does not believe that there are any existing general things.⁴⁵ As discussed earlier, Locke's epistemology distinguishes between ideas from sensation and reflection. Since there are no existing general things to allow the idea to come from sensation, the idea of abstract ideas must be from reflection.⁴⁶ The force of Berkeley's and, subsequently, Hume's attacks is usually taken to be their proof that there cannot be any existing general things. However, since Locke never claimed that any such things existed (he even also explicitly denies the possibility of their existence!), their arguments were merely of the *straw man* nature.

Earlier, it was implied that there may not be any abstract ideas. In order to address and resolve such a statement, one must first define what an abstract idea is. One can either take Locke's conception, which includes the perfect separability of qualities, or Hume's «distinction of reason.» Recall that Hume's distinction of reason is not merely our ability to distinguish between different ideas in our mind that are really inseparable in the objects themselves; instead, it is an *illusion* that we are making this separation in our minds: *strictly speaking*, we are not able to separate the ideas even in our minds. Since it is truly impossible to think of a colour without some extended body on which to imagine it, I find Locke's belief in the separability of colour from extension impossible. Thus, I find his conception of abstract ideas to be untenable. Hume's account, however, seems to fix this – through his distinction of reason – and has a more complete account of what abstract ideas are than Berkeley (by discussing our ability to use a distinction of reason to form a different sort of abstract idea than that of Locke's). The only point of departure between Hume and Locke, with respect to abstract ideas, is that Hume does not believe in the perfect separability of qualities (separated in the mind). Consequently, there seem to be abstract ideas under Hume's conception, but not Locke's.⁴⁷

In summation, a close reading of Locke's actual arguments has shown that not only have many of Berkeley and Hume's arguments been of a *straw man* nature, but, other than the one (major) point of departure (the inseparability of qualities), they seem to *agree* with Locke's account of the genealogy of abstract ideas, and their corresponding general terms (the terms are merely signs for the ideas).⁴⁸ Unfortunately, due to space constraints, this paper was not able to discuss the full extent of either Berkeley's or Hume's accounts on the genealogy of abstract ideas (Hume's argument from «custom» for example), but I do believe that the issues surrounding the *actual* strength of the attacks on Locke by these two prominent philosophers have been furthered.

⁴⁵ Essay, 3.3.11. See note 10 as well.

⁴⁶ By using the disjunctive syllogism (process of elimination), see note 20.

⁴⁷ Since I found Locke's definition of abstract ideas, *viz.* including the perfect separability of qualities, to be untenable.

⁴⁸ I say «they» because Hume is, in his mind, arguing in favour of Berkeley; however, Berkeley may have also rejected Hume's account of abstract ideas *qua* distinction of reason (but this is only conjecture since Berkeley never *explicitly* brings it up).

Works Cited:

1. Dicker, G., *Hume's Epistemology and Metaphysics*. (London: Routledge, 1998).
2. McKeon, R., ed. *The Basic Works of Aristotle*. (New York: Random House, 2001).
3. Nidditch, P.H., ed. *John Locke: An Essay Concerning Human Understanding* (Oxford: Clarendon Press, 1975).
4. Norton, D.F. & Norton, M.J., ed. *David Hume: A Treatise of Human Nature*. (Oxford: Oxford University Press, 2002).
5. Woolhouse, R., ed. *George Berkeley: Principles of Human Knowledge/Three Dialogues* (London: Penguin Books, 1988).

Rhys McKinnon
University of Victoria
British Columbia, Canada
<rhys.mck@gmail.com>

TARSKIAN METAMATHEMATICS IN CARNAP'S METALOGIC

Jesús Padilla Gálvez

Introduction

An important group of logicians working in the field of metatheory proposed a variety of different suggestions on the description and analysis of signs of formal structures. Thereby they influenced each other in different ways. This metatheory was then called «metamathematics», «metalogic», or «methodology of deductive sciences». Divergence certainly did not rely on the same mode for all logical significant aspects of a system, but with regard to the fundamental frame of references and composition of basic elements of the structure of formal systems, which should be described and analyzed. The aim of this paper is to view the contours of metalogic through a careful and exact analysis of the works of A. Tarski and R. Carnap and to differentiate between the various approaches, formulations and interests. In this way their differing positions on formalistic problems, divergence and convergence can be shown in full. This task demands a wide-ranging and exact revision, in the course of which it is easy to lose sight of the basic logical structure. The subject of this paper is to attempt a systematic explanation of metalogic.

1. Historical approach

Is it not surprising that one of the most influential members of the Vienna Circle, R. Carnap, supported metalogic despite his «metaphysical suspicions» and introduced the idea of looking at the problem of logical syntax of language from precisely this point of view? How could metalogic comply with the logistical basics of mathematics? How could different directions and philosophical points of view be connected to each other, without denying their positivistic aspect? This and other questions point to the problems which were discussed in the Vienna Circle in the Thirties.

In 1931 Carnap did not share Gödel's arithmetic view but held a descriptive view. The question which interests us is whether Tarski's descriptive view of metalogic was accepted. Carnap later changed his views in his «Syntax» and accepted and integrated Gödel's pioneering results into his program.

On 19th February 1930 Tarski gave a talk in Vienna with the title: 'Über das Auswahlaxiom und die verallgemeinerte Kontinuumshypothese'. On the following two days Tarski spoke in the Schlick Circle on 'Metamathematik und Metamathematik des Aussagenkalküls'.¹ The following question was up for debate: under what conditions can we speak of formal languages in a legitimate way free from objections, without hitting upon the

¹ A. Tarski (1992): «Alfred Tarski: Drei Briefe an Otto Neurath» (ed. by Rudolf Haller. Translated into English by Jan Tarski), *Grazer Philosophische Studien*, 43, 1-31. S. 23.

common difficulties of self-references? The two great dangers of logical research were firstly those of autonomous ways of speaking² and secondly, drawing attention to the fact that expressions which we «indicate» have a relative character and therefore must always be relativised in a particular language. Metamathematics advocates a clear distinction between logic-mathematical formalism (the so-called object language) and the metamathematical considerations. This is particularly expressed by the difference between symbols and variables of formalism and signs of «communication». The formal theory is therefore opposed to a metalogic or metamathematics. Tarski attempted to show that although questions can be systematized by the expansion of expression in certain languages compared to the calculated language the systematic nature of which can be dealt with in an axiomatic form.

Both speeches divided the Vienna Circle. Carnap wrote in his diary that Schlick did not recognize the relevance of the metamathematical investigations.³ It is possible that Neurath's resistance to Tarski's views developed during the lectures. At first it was suspected that Tarski's metalogical definitions (in particular the definition of truth) were not without requirements. The general mistrust within the Vienna Circle was not the same amongst all the members. In a frequently quoted passage from Carnap's letter to Neurath on 23rd of December 1933, he gave a summary showing which root his «syntax» had originated from. He wrote: «My syntax has historically two roots: 1. Wittgenstein, 2. Metamathematics (Tarski, Gödel).»⁴ This reference is even more surprising as firstly, Wittgenstein had always been critical of metalogic and secondly, in the area of metamathematics, the Hilbert school was not mentioned. The meaning behind the Wittgensteinian reference can only be related to the fact that he had drawn attention to the importance of the problems which affect the language. In his research, however, Wittgenstein always rejected the view that it would be legitimate to talk about language and this argument took hold amongst some of the members. The second aspect, that the Hilbert school had not mentioned, requires an investigation of its own.

It is not our intention, however, to dwell on the negative results and discussions, rather it should be to emphasize the positive aspects. It is therefore appropriate to raise the following question: What positive effect did Tarski's lectures and discussions bring about? Was Gödel's work really so fundamental to Carnap's program in metalogic? In the secondary literature, opinions are divided. Alberto Coffa and others believe that Gödel's results, published in 1931, were the most important influence on Carnap.⁵ Most academics do not agree about Carnap's critical position towards Gödel. In contrast to other interpretations, some authors believe that Tarski's Vienna speeches in February 1930 gave great impetus to Carnap's syntactical conception of language. This impetus is necessary for the so called «Erläuterungssprache» (i.e. explanatory language, language of elucidation or metalanguage) to be symbolized exactly.⁶

² According to the autonomous ways of speaking «... should constantly distinguish the sign from the designated object, especially in the cases, where the designated object is again a sign or, more generally, a figure of speech.» Tarski (1992), 26.

³ Tarski (1992), 4.

⁴ Carnap's letter to Neurath 23 December 1933. Hilman-Library, RC 29-03-06 A.

⁵ A. Coffa (1991): *The Semantic Tradition from Kant to Carnap: To the Vienna Station*. Cambridge.

⁶ R. Carnap's 'Tagebuch'. See: Tarski (1992), 5.

It would be very simple in this view to indicate that the «Erläuterungssprache», is the language that we need in order to talk about the object language. Tarski spoke out very clearly against this simplification when he wrote: «the simple statement that sentences about sentences are legitimate, appears to be completely unfounded.»⁷ In discussions a dichotomy between the «Erläuterungssprache» and metalanguage is indicated and therefore an important part of the discussion is blurred. To adopt a completely different approach, I would like to analyze the discussion and in so doing, explain that the differences in the ‘Erläuterungssprache’ can be shown and that differences within the research areas between the Tarskian and the Carnapian programs can be highlighted. The question I would like to examine is this: what method did Tarski use to conceptualize his «Erläuterungssprache». Or to put it another way, how did the need for an «Erläuterungssprache» originate?

2. From the formal system to ‘Erläuterungssprache’

Yet can we ask ourselves what they understood then by the so-called «Erläuterungssprache» and how they reached that definition? In order to answer the questions in detail, we can see that it is necessary to look for sources and this is where we encounter difficulties. In the «Erläuterungssprache» no reference is made to meaning. If one looks at the meaning of a language, one can analyze the language mathematically. The resulting system contains both the description or definition as well as the study of the qualities of formal systems. In an abstract way, Tarski both describes and defines the deductive systems⁸ in a narrower and more general way than usual.⁹ On the one hand the definition is narrower than normal as it touches on the definition of logical consequence which is defined in classical logic. On the other hand, it is more general as it refers to naive viewpoints of set theory. With regard to the second definition, a system is a set of propositions. For Tarski it was important to differentiate between the invariable deductive systems and the consequence relation. Tarski referred to the closure of all consequences.

The naive view of the finiteness of a set goes back to the well-known concept of the natural number. The starting point of a theory of finite sets forms a definition of the concept of the finite set. Depending on the choice of these starting definitions, the theories will develop differently and individually. Tarski’s definition has the advantage that on the decision about the finiteness of a set, only this itself and its subsets, but also further sets of a general character do not need to be referred to.

In order to study the question of logical consequence in the correct structure, it is necessary for us to briefly look at the definition of formal systems. A formal system is a hierarchical grouping of sets of symbols or complete formulae, out of which other formulae can be generated, which can be accepted as valid. So that a complete system of interests can be created, a kind of lock must be made available, which can be viewed from two different and independent sides. On the one hand, not every complete formula of the system can be

⁷ Tarski (1992), 26.

⁸ A. Tarski (1935) «Der Wahrheitsbegriff in den formalisierten Sprachen», *Studia Philosophica*, 1, pp. 261-405/ 503-526; A. Tarski (1935-36) «Grundzüge des Systemkalkül», Erster Teil, *Fundamenta Mathematicae*, 25 (1935), 503- 526; Zweiter Teil, *Fundamenta Mathematicae*, 26 (1936), 283-301; and A. Tarski (1936) «Über den Begriff der logischen Folgerung», *Actes du Congrès International de Philosophie Scientifique*, Paris 1935, Vol. VII, ASI, 394, Paris, 1-11.

⁹ Tarski (1992), 27.

proved; on the other hand, though, it is impossible to constantly add new complete unprovable formulas to be produced, without every complete formula being provable. The precise definition of both qualities depends on the deduction rules which fixes the concept of proof.

By *propositional statement* or *meaningful sentence* we understand «... certain inscriptions of a well-defined form».¹⁰ Somewhat later he published his thoughts which he had already written in 1931 and can be characterized in one sentence «... as a particular kind of expression, and thus as linguistic entities.»¹¹ From this a definition was drawn up to make clear the difference between «use» and «mention». This difference leads to the assumption that the statement does not denote concrete series of signs but the whole class of such series which are of like shape with the series given. After this, the quotation of names could be treated as individual names of expressions.

After his return from Poland, Carnap delivered three papers entitled 'Metalogic'. In metalogic he saw the focus of his research as being the analysis of the quotation marks of a particular language.¹² The first priority of Metalogic is to answer (i) which signs appear and (ii) which row of symbols are formulae. A metalogical sentence serves as a description. He understood «description» as the presentation of «empirical data» which comes from «... Belegung eines Stellegebietes mit Qualitäten (oder Zustandsgrößen)» besteht.»¹³ In Carnap's view it seemed that a formula is seen as an elementary disjunction from the metalogical definition of a concept of «elementary disjunction», which has a metalogical description.

Four years later in the 'The Logical Syntax of Language' a sentence was defined as an expression which corresponds to a statement of natural language.¹⁴ In this way a distinction was made between Language I and II. Language I only contains definite concepts, whereas Language II is much richer in terms of expression. Furthermore it also contains indefinite concepts, has classical mathematics and can also formulate sentences of Physics.¹⁵ In the field of metalogic a language is a «...sort of calculus ..., a system of formation and transformation rules concerning what are called expressions, i.e. finite, ordered series of elements of any kind, namely, what are called symbols.»¹⁶

Tarski and Carnap differ in their general views on what they understand as metamathematics or metalogic. For Tarski, metamathematics is a «General theory of

¹⁰ A. Tarski (1930) «Über einige fundamentale Begriffe der Metamathematik», *C. R. des Séances de la Société des Sciences et des Lettres de Varsovie*, Cl. III, 23, 22-29, see p. 23.

¹¹ Tarski (1935), 269; note 5.

¹² R. Carnap (1995) *Metalógica-Metalogik* [1931], *Mathesis*, 11, 137-192; see p. 139.

¹³ Carnap (1995) [1931], 140.

¹⁴ R. Carnap (1934) *Logische Syntax der Sprache*, Wien-New York, Springer Verlag; p. 13.; R. Carnap (1937) *The Logical Syntax of Language*, London, Routledge & Kegan Paul Ltd; p. 14.

¹⁵ Carnap (1934), 78/ (1937), 89.

¹⁶ Carnap (1934), 120. (1937), 167f.

mathematical sciences». He later showed very clearly that the strengths of the Warsaw School had been used in his metamathematical research. He himself said: «As an essential contribution of the Polish School to the development of metamathematics one can regard the fact that from the very beginning it admitted into metamathematical research all fruitful methods, whether finitary or not.»¹⁷ It examined the «... mathematical theories in their entirety».¹⁸ For Carnap, however, metalogic is a «... Theorie der Formen, die in einer Sprache auftreten, also die Darstellung der Syntax der Sprache.»¹⁹ It is «... the theory of the forms of the expressions of a language».²⁰ In order to reveal the main similarities and differences, we see that we need to look much more closely at the concept of consequence. An exact definition of the concept can be given in metamathematics or metalogic where the research subject matter forms a solid formalized study. Tarski sees the concept of logical consequence as «a primitive concept» and characterized it by means of a number of axioms which we shall now look at in more detail.

3. The notion of «Logical Consequence»

Consequence always leads to new knowledge without the help of experience. If it is known for instance that Vienna lies to the east of Linz and Linz to the east of Salzburg then one can conclude from both these statements that Vienna is east of Salzburg. Every deductive conclusion is accordingly an example of this consequence. When it is known that judgements contain subject and predicate, and likewise that a proposition is also a judgement then it can be concluded from both premises that this proposition contains subject and predicate similarly. A logical consequence in the rules governing conclusions of relations and connections between the genus and the type is displayed. As that which makes a genus unique also makes its species unique, hence one must copy the model in the operative relations between genus and concepts of species.

Earlier Tarski had drawn attention to the evidence of a given set underpinned by certain operative functions, from which the consequences of the set of evidence can be constructed.²¹ Nevertheless Tarski did not follow this up until later. This prompted Carnap to pick up the theme in the course of his investigations into the difference between the syntactical elements of derivation and the semantic elements of deduction.²² Viewed systematically the concept of «consequence» should be dealt with at the outset of metalogic. If the concept of consequence specific to a language is determined, the logical connections within the language are laid down.

¹⁷ A. Tarski (1986) *Collected Papers* (Eds. St. R. Givant, R. N. McKenzie). Vol. 4. Basel, Birkhäuser. p. 713.

¹⁸ Tarski (1995) [1939], 159.

¹⁹ Carnap (1995) [1931], 139.

²⁰ R. Carnap (1963): «Intellectual Autobiography». In: P.A. Schilpp (Ed.), *The Philosophy of Rudolf Carnap*. La Salle, Ill. Open Court, 1-84; see: p. 54.

²¹ Tarski (1930), 97ff. and Tarski (1935-36).

²² Carnap (1934), 88ff.; 124ff.; 128. / (1937), 98ff.; 170ff.; 175.

Carnap distinguishes two types of deductive processes, namely derivation (Ableitung) and consequence (Folgerung).²³ In the syntactic derivation d-terms are to be found: On the one hand, derivable, demonstrable, refutable, resolvable and irresolvable. On the other hand, there are the semantic consequences belonging to the family of the c-terms: consequence, analytic, contradictory, L-determinate and synthetic.²⁴ In Language I is used the concept of «consequence», to define notions of «analytical» and «contradictory». Every sentence and every class is either analytical or contradictory.²⁵ By analytical he understands a sentence involving a consequence of the null class of sentence and accordingly the consequence of every sentence. Every demonstrable sentence is analytical. The concept «analytical» refers to what is demonstrated to be logically valid or true for logical reasons. A sentence is regarded as contradictory when every sentence is the consequence of the null set. This also applies to a sentence class, when every sentence is the consequence of a negative sentence class. Every refutable sentence is contradictory. But the sentence class is only contradictory when at least one sentence belonging to it is contradictory.

In contrast the reverse process is suggested in a simplified form for the Language II. Furthermore the concepts analytical and contradictory are defined for sentences and sentence classes and by extension for the concept of consequence. Carnap's concept of an analytical sentence is dependent on every given class having an analytical quality or on the syntactic quality of expression, the analytical sentence which can be demonstrated in the following way from this sentence: every part of a sentence is constructed in the case of the sentence part belonging to the class through the $0=0$ ²⁶ and otherwise through the negation of the $0=0$.²⁷ Tarski referred to Carnap's first attempt also formulating a precise definition of the concept of consequence. He refuted Carnap's attempt, because his suggestions depended on special qualities of the formal language. According to Tarski Carnap's position runs as follows:

(FC) «The sentence X follows logically from the sentences of the class K, if and only if, the class consisting of all the sentences of K and of the negations of X is contradictory.»²⁸

Tarski attacked Carnap for shifting attention away from the concept of logical consequence towards the concept of the contradictory. The definition was complicated and highly specific. As an alternative, Tarski attempted to maintain an essential condition, namely statement X from a class of statements K following which he described thus: (FT) «If, in the sentences of the class K and in the sentence X, the constants - apart from purely logical constants - are replaced by any other constants (like signs being everywhere replaced by like signs), and if we denote the class of sentences thus obtained from K by 'K'', and the sentence

²³ Carnap (1934), 36 / (1937), 41f.

²⁴ Carnap (1934), 88 / (1937), 101.

²⁵ Carnap (1934), 34ff. / (1937), 37ff.

²⁶ Carnap (1934), 75 / (1937), 84.

²⁷ Carnap (1934), 88 / (1937), 102.

²⁸ Tarski (1936), 6.

obtained from X by 'X', then the sentence X' must be true provided only that all sentences of the class K' are true.»²⁹

Given that the condition (FT) is sufficient, then the appropriate definition of the concept of consequence would be positively decided. Certainly a proviso would have to be attached to the definition, namely the use of the semantic concept «true».³⁰ The use of the semantic concept of the predicate «true» is the essential difference to Carnap's definition of the concept of consequence which is founded on the concept of contradiction.

4. The peculiarities of analytic arguments

Why did Tarski refute Carnap's proposition about logical consequence? Tarski refuted Carnap because the latter stated in the supposition that certain objective problems can be reduced to linguistic problems. This reduction led to a false interpretation of the a priori point of view and of the role it plays in the analysis of reality. This opinion was taken of by Wittgenstein, who claimed that all a priori propositions, i.e. those that belong to logic and mathematics are of a tautological nature. Carnap however used other terminology and called all these propositions «analytic». Sentences which strictly adhere to the convention and which do not make material modes of speech in a statement are known as analytic.

All supposed a priori investigations are analytic, in the same way as conventionally accepted rules which govern the use of certain expressions in our language. Of course these interpretations of language are supported by Wittgenstein's interpretation which stated that all languages have the same logic. Therefore every analytical sentence in the context of Wittgenstein's interpretation of language, independent of any empirical proof, is correct. Thus the analytical concept can assume the role which in traditional philosophy is known as «a priori». In any case, the concept of analytic seemed to be a satisfactory representation of logic and mathematics. The problem arose when Wittgenstein's interpretation of language began to waver due to the development of an 'Erläuterungssprache' on the metalevel.

Tarski moved away from the terminology of the earlier Vienna Circle because it had accepted an association which could have led to a misinterpretation. Both Wittgenstein and Carnap held that a priori propositions say nothing about reality, as they are simply instruments which make the recognition of reality possible. If necessary a scientific interpretation of the world can be given without having to refer to a priori elements. Tarski started with another assumption as he was very conscious that even since the «Grundlagenstreit», alternative logical systems were seen as evolving separately from each other. Furthermore they were thought to possess the quality that one can be translated into the other. Even today we cannot decide whether the relations between facts are best demonstrated through classical logic or through polyvalent logic's.

The great debate between the members of the Warsaw and the Vienna Circles was therefore the question of what effect Wittgenstein's conception of language would have. The Grundlagenstreit about logic had led to a new scenario in philosophy because the expressions of non-equivalent languages (such as those of logicism, of intuitionism and of axiomaticism) cannot be used interchangeably. As a consequence the supposition of a language- neutral

²⁹ Tarski (1936), 7.

³⁰ Tarski (1935), 261-405.

concept of analytic was dismissed. Analeptics became (as Tarski had always implied) a language related concept.

Neither Carnap nor the Vienna Circle were prepared for this situation. It should be noted that in 1931 Carnap had recognized the relevance of the *Erläuterungssprache* for a formal system, but he did not perceive the philosophical consequences that this entailed. One consequence that we have looked at is the revelation of the unity of the language of sciences. Carnap also did not recognize all of the implications of this in 1935.³¹ Metamathematical investigations, i.e., according to the Warsaw school, can then be carried out if the concept of statement and consequence is precise. Later in 1930 a program was devised which as its starting point had a set statement and was analyzed in relation to the lack of contradiction and axiomatic. The level of completion would then be tested.³² In this program the work was carried out exactly as laid down in Hilbert's *Basics of Arithmetic's*: the so-called «geometric change»³³ was completed. Tarski began to interest himself in the supposition of the concepts «statement» and «consequence» and brought together these elements using certain axioms. This brought about the necessary consequence of showing the weakness of contradiction and axiomatic. Finally he tried to reach a more complete analysis of axioms. In this way Tarski came to accept the elementary-arithmetical character of metalogic.

Jesús Padilla Gálvez

Universidad de Castilla - La Mancha
Facultad de Ciencias Jurídicas y Sociales
E-45071 Toledo
<jesus.padilla@uclm.es>

³¹ S. C. Kleene (1971): *Introduction to Metamathematics*. Amsterdam, North- Holland Publishing, p. 65.

³² Tarski (1930), 28 f. A. Heyting (1930) «Die formalen Regeln der intuitionistischen Logik», *Sitzungsberichte der Preussischen Akademie der Wissenschaften*. Physikalisch-Mathematische Klasse II, 42-56. K. Gödel (1930) «Einige metamathematische Resultate über Entscheidungsdefinitheit und Widerspruchsfreiheit», *Anzeiger der Akademie der Wissenschaften in Wien*, mathematisch-naturwissenschaftliche Klasse, 67, 214-215.

³³ P. Bernays (1923) «Erwiderung auf die Note von Herrn Aloys Müller: 'Über Zahlen als Zeichen'». *Mathematische Annalen*, 90, 159-163.

WHY AXIOMATIZE ARITHMETIC?

Charles Sayward

P. The axioms for arithmetic, and deduction from them, are epistemologically irrelevant.

For it is necessary, if we are to be justified in accepting this or that as an axiom, that its logical consequences not conflict with what is already known about the numbers, but instead include things already known about the numbers.

So we cannot be justified in accepting the axioms unless we *already* know a lot about the numbers *independently* of the axioms or any deductions from them.

After all, its not as if mathematical knowledge *began* with Peano!

C. Your idea seems to be this: Lots of things get to be known about the numbers. Then someone proposes certain sentences as axioms. So the question arises, ought we to accept these sentences? And we decide whether we should by seeing (i) whether they logically entail antecedently known arithmetical truths, and (ii) whether they logically entail anything inconsistent with antecedently known arithmetical truths. Only if they entail what we already know, and entail nothing which conflicts with what we already know, are they acceptable axioms.

Isn't this your thought?

P. Yes.

M. But the statements we accept *as axioms* were among the ones *already* accepted *as true*.

Peano's proposal was not that certain sentences (i)*are true* and (ii) *may serve as axioms*. Rather, he selected certain already known arithmetical truths and proposed only that they may serve as axioms.

Peano was not the first person to suppose that zero is the first of the natural numbers or that different natural numbers have different successors! His innovation lay not in *asserting* the things his axioms assert — as mathematical assertions his axioms were already familiar and accepted — but in proposing that these statements were *logically sufficient* for the whole of the arithmetic of the natural numbers.

P. Fair enough. But my point still holds good: we can be confident of the *truth* of the axioms only if their consequences, as so far known, agree with what we *already* know about the numbers, both in the sense of conflicting with nothing already known and actually leading to things already known.

For suppose someone were to propose as an axiom some sentence which logically entails something which goes against some known mathematical truth. Wouldn't we then reject the proposed axiom?

C. I find the question virtually meaningless.

Suppose that some particular equality — for example that $5+7=12$ — logically entailed something which went against some piece of mathematical knowledge already in our possession. How would you answer the question about what we would then reject?

Wouldn't you rather become suspicious of the question, or the supposition upon which it rested?

P. I don't follow you.

C. Well, we are supposed to suppose that we discover that the statement that $5+7=12$ logically entails something which conflicts with something we already know. So suppose, if you can, that it logically entails that $4=5$.

Now, what gives way? Do we reject the equality $5+7=12$ or the inequality $4\neq 5$?

Suppose we reject the inequality. Then we agree that $4=5$. But $4+0=4$ and $4+1=5$. Thus, since $4=5$, $1=0$. But then, since 2 comes right after 1, 2 comes right after 0. But 1 comes right after 0. So, $1=2$. So, $0=1=2$. And thus also $0=1=2=3=4=5=6\dots$ That is, then all the numbers are one and the same!

Well — that makes it look as if we'd better reject the equality. Then we must hold that $5+7\neq 12$. What then is $5+7$? Well, it must be greater than or equal to 7. Suppose it is 7. Then, $5+7=7$, in which case $5+7=0+7$. But then $5=0$. But since 6 comes right after 5, 6 comes right after 1. So $6=2$. By the same reasoning $7=3$ and $8=4$ and $10=5$ and $11=1$, and so on. That is, *all* the numbers are just the numbers 0 through 4. Now suppose that $5+7=8$. Then $5+7=1+7$, in which case $5=1$. But then $4=0$, in which case there are just the numbers 0 through 3. And so on.

So the situation is disastrous on *both* alternatives.

P. I see that.

M. The same sort of thing arises in respect to the «supposition» that some axiom entails the equality that $4=5$, for example the axiom that zero is the first number. What would we reject if it turned out that if zero is the first number then $4=5$? If we reject the axiom, then we agree that *some* number comes before zero. But then zero comes immediately *after* some number. So let the number be one. Then zero comes right after one. But so does two. But no *two* numbers come right after any number. So, we conclude that $0=2$. And if that is so, then $0=3=4=5=6=7=8=9\dots$ just as well. And if it isn't *one* that zero comes after, it is *some* number — and the conclusion will be that the numbers then stop with *its* immediate successor, namely zero!

So, if we reject the axiom we will have to accept something outlandish, such as $0=2$. If we do not reject the axiom we will have to accept something equally outlandish, for example that $4=5$.

So, and please do not take this remark badly, it seems to me that no *thinking* has gone into this idea that we must test the axioms against what is already known. It has just been words without thought.

P. Perhaps there is something in what you say. I do now feel as if I never bothered to think through the supposed supposition I was making. But what if someone proposed something *brand* new as an axiom — then, surely, it would have to meet the test of agreeing with what was already known in arithmetic, and so couldn't be a *source* of knowledge.

C. But what is this «something brand new»? Presumably it will be a sentence of arithmetic. So it won't be *syntactically* or *semantically* novel. So it must be one whose *truth-value* is as yet unknown to us.

But *no one* proposes as an axiom for arithmetic some arithmetical sentence we aren't already sure of!

P. But surely this *could* be done. Didn't Gödel suggest that we might try to think up further axioms for set theory? And any such axiom would not really be *known* to be true at the start?

C. Perhaps. But right now we are talking about axiomatic arithmetic, not axiomatic set theory. The two may be quite different, and certainly the intellectual pressures that gave rise to them are entirely distinct. So let's just stick to arithmetic. And here my point is this: In arithmetic no one proposes as axioms statements about which we are as yet *uncertain*.

P. As I reflect on what you say, I find myself forced to agree. The actually proposed axioms, and very likely any that anyone would seriously propose, are statements about whose truth we are already satisfied. In a word, and this is the point I did not appreciate, axioms are selected from what is *already known*.

I think that what happened in my thinking was that I slipped from asking what would be needed to accept a statement *as an axiom* to asking what would be needed to accept the statement. Then, since a condition on its being an axiom is that a statement, together with the other axioms, leads to all the already recognized truths, I slipped into thinking that this was a condition on accepting the statement itself — as if it needed to pass a test taken simply as a statement of arithmetic!

In any case, I now seem to see that it is quite correct to say that the statements which are proposed as axioms for arithmetic need pass no special test since they are all among the statements already known to say true things about the numbers.

But though I will grant that proposed axioms state about the numbers things already known, surely there is much in addition to what is said by those axioms that also is *already known* about the numbers. And those further, *already known* things are not known by *deduction* from the axioms. That point remains secure.

And so I will continue to assert that the axioms and deductions from them are epistemically irrelevant.

C. But if other things in arithmetic *could* be known by deduction from the axioms, then if we can see how we *do* know that the numbers are as the axioms state them to be, then we can see how we *can* know all the rest. We could gain a certain insight into the *knowability* of the propositions of arithmetic, even if not into the manner in which all or some of them have actually become known.

P. But it well may be that we couldn't know that the numbers are as the axioms state them to be *without* already knowing ever so many of the more specific points about particular numbers which follow from the axioms. And in that case we will not gain the insight to which you refer.

I will make my point this way: You assume that the axioms can be known *on their own* and apart from knowing anything else pertaining to the numbers. But this assumption may be in error.

C. Is that my assumption? It may be a kind of psychological impossibility to know that zero comes before all the other numbers and *not* know that it comes before one. For it might be impossible for us *not* to see that zero comes before one if it comes before all numbers other than itself. If so, we could not recognize that zero comes first without recognizing as well that zero comes before one. So it might be impossible to know the truth of the general proposition without knowing the truth of some of its instances. I have not suggested or assumed the opposite.

But that would not show that knowledge of the truth of the general proposition is *based* on knowledge of the truth of its instances.

P. On what else might it be based? Can't we see that it *must* be based on knowledge of the truth of at least some of its instances?

C. I am not at all sure on what our knowledge of the truth of the axioms is based. But I think I can see that it is *not* based on such knowledge of the truth of its instances as we may possess. And here is how I see the matter. Consider the axiom which states that zero precedes every other number. Its instances are 'zero precedes one', 'zero precedes two' and all the rest. Now, we know the truth of only finitely many of these instances. But we surely couldn't know that zero precedes *every* other number by knowing that it precedes three or fifty or fifty thousand other numbers. So our knowledge of the truth of the axiom cannot be based on our knowledge of the truth of its instances.

If we came to believe that zero precedes every other number by noting that it precedes one, and two, and three... and five million, then that would be merely a *conjecture* — not a piece of knowledge. And a very weak conjecture at that, since we would be extrapolating from the finitely few to the infinitely many.

Yet, we surely *do* know that zero precedes every number other than itself. And so it seems that this knowledge *is* independent of such knowledge of the truth of its instances as we may possess. And the same holds for each of the other axioms. So they all are known *independently* of their instances. But their instances can be deduced, and thus known through them.

So that is *one* way in which knowledge is possible in arithmetic, and *this* way of knowing will become fully clear to us as soon as we make it clear to ourselves how we *do* know the truth of the axioms.

P. I feel very uncomfortable about this. It continues to seem clear to me that we know e.g., that zero does not come after one, and know that every bit as much as we know that zero follows no number and do *not* know it by deducing it from that general proposition.

C. But I have said only that by reference to the axioms we can arrive at an understanding of how arithmetical knowledge *can* be obtained — not that we will arrive at an understanding of how arithmetical knowledge *must* be obtained, or even *has* been obtained.

It may be that we *can* know that zero comes before thirteen by deducing that proposition from the known truth that zero comes before every other number, and that is consistent with our *actually* knowing that zero comes before thirteen *independently* of any such deduction.

I am not suggesting that we ordinarily *do* know, much less than we *must* know, the familiar sums and products *by deduction* from *independently known* axioms. I say only that we *can* know these sums and products by deduction from independently known axioms.

P. I see what you're saying, and I don't see how to object to it. Yet there seems to me to be *something* wrong in your conception of the matter.

It seems so clear to me that knowledge of the truth of various instances *comes first* and that knowledge of the truth of the generalization *comes second* and *couldn't* come first. And so I feel that knowledge of the truth of the axiom really is secondary or derived. Despite your arguments, I cannot shake myself of this feeling.

C. Well — you *say* this, but then you don't explain it. So I am quite at sea about how best to respond to you. I really don't see why you say we *couldn't* know the truth of some consequence of an axiom by deducing it. Deduction surely is *one* way of obtaining further knowledge from some already obtained knowledge. I hope you will not go so far as to question *that*.

P. That has tempted me. But let me try this instead.

As I understand it, people had a good grip on a lot of arithmetic *prior to* getting a handle on zero. This shows that people could know — *did* know — enormously much about the numbers *without* so much as *conceiving* the content of most of the axioms — for most of them include the idea of zero.

C. Agreed. But that would be a somewhat different system of arithmetic, and so would have somewhat different axioms. Instead of asserting that zero comes first among the numbers, the right axiom for this arithmetic would assert that one comes first among the numbers.

But what turns on your point? I have not denied that our actual knowledge may be independent of deduction. I have only asserted that deduction from antecedently known axioms is *a* way of knowing.

P. But the important point is that it is possible to have *fragmentary* arithmetical knowledge. And just as there is arithmetic without zero, there can be arithmetic without generalization. People can know a lot about individual numbers — their sums and products — without as yet so much as *considering* generalizations like 'Zero comes after *no* number'. It might not be the case that anyone who *considers* that statement fails to recognize its truth. But people might be, as it were, oblivious to it and *still know lots of things about lots of numbers*.

So I will now put my point this way: Our knowledge of the individual numbers — as contrasted with our knowledge of arithmetical laws — in *no* way depends upon our knowledge of those laws.

C. But now you are no longer viewing the epistemic irrelevance of axiomatic arithmetic in the same way that you did at first.

P. I agree. My initial «view» made no sense at all. It was, as you put it, words without thought.

But that recognition did not free me of a feeling that somehow the laws of arithmetic are *secondary*. And so, I am now working on *that* idea.

What I now wish to say is that the fact that we can know enormously much about the numbers without so much as considering the laws shows the irrelevance of those laws to our basic arithmetical knowledge.

C. In a way what you say seems undeniably correct. But aren't you still missing my main point — for I have only said that it is *possible* to arrive at arithmetical knowledge by deduction from axioms independently known to be true. Not that we must or even actually do so arrive at arithmetical knowledge.

P. Yes. For some reason I keep slipping up on that point. Still — there is something I am after, so let me shift ground again.

What I now want to say is that knowledge of an axiom *does* depend on knowledge of its instances. Not all of them, but some of them. We first come to know instances. We perceive a pattern in the instances. We then generalize on the known instances to bring out that pattern. I do not mind calling that a kind of speculation. Then by deduction we bring out further particular points that we can then put to the test.

This, I now think, is how things work in arithmetic. We know the axioms to be true only by seeing that again and again their entailments are true.

C. Here you're viewing the axioms as analogous to expressions of empirical regularity. This much is right in that analogy: the axioms are laws. They are laws of numbers.

You're also suggesting that the laws of number are hypotheses suggested by particular facts and then tested by seeing what further particular facts they lead to.

This makes knowledge of the laws entirely secondary.

But I feel that *in some way* knowledge of laws is present in even the most elementary pieces of arithmetical knowledge — in a way in which knowledge of an empirical generality is not involved in knowledge of its instances.

I could put how I feel about the matter this way: I need not in any way work with the idea that all ravens are black to spot the blackness of *this* raven.

So I just cannot accept the sharp distinction you draw between knowledge of the laws of number and knowledge of the numbers.

P. That is extremely obscure.

C. I agree. And I could say even more obscure things than that on the topic! But don't worry about the obscurity, for even an obscure remark might lead to better things.

So let me go on and try this assertion, that there is something *general* involved in what goes into knowing anything arithmetical.

P. I would say that you were true to your word, you can indeed increase the obscurity of your pronouncements!

But seriously, tell me what you have in mind — if you *do* have anything in mind.

C. Well, something general comes into learning the numbers — learning which numbers are which and what the sums and products are. For you don't just learn this number and then that one, in a random fashion, as if you were learning about this raven and then about that one, but you learn a *system*. You learn to go from one to two, and from two to three, *and so on*. You need to do something like master a rule. And there is something general about a rule.

I would not deny that this is still very obscure. But it feels right nonetheless, and perhaps just a bit less obscure than what I began with.

P. But don't some people count one, two, many? Or a child might master just the numbers through five.

C. Yes. There are these fragments of arithmetic which, taken on their own, are in some ways like and in more ways unlike arithmetic. But if you've got the *numbers* then you don't have just five or fifty of them. In a certain sense you have them *all*. And that's where the generality or law comes in. We might say that you master a law of construction for the numbers, or for recognizing them. And it is the mastery of *that* law which gets reflected in the first two axioms.

A person might be thrown by the question 'And is there a greatest number?' while yet knowing full well that if there is one more swan than fifty there are fifty one swans, and if there is one more swan than fifty thousand, then there are fifty thousand and one swans. *And so forth*.

And the person thrown by the statement that one comes before every number other than itself will nonetheless always start with one when counting! Here, we might say, a certain feature of his or her practice constitutes a kind of recognition that one comes before all the other numbers — and it is this feature of *practice* that is explicitly set out in the axiom.

Even a child will not be said to as yet know its numbers if its counting-like actions begin at times with one, and at times with three, and so on. Counting isn't just any kind of repetition of familiar sounds together with gestures of pointing.

P. But surely miscounting is counting. Just as invalidly inferring is inferring.

C. A child can miscount only if it has mastered the technique of counting. The very acts that with us are miscountings would not be that if they constituted the normal or typical performance of a child.

But we are drifting. Let's review where we're at.

We've moved from talking about axioms and their instances to the topic of the general and the particular in arithmetic. The question we are now considering could be put as follows: Is knowledge of the laws of number somehow *implicit* in our knowledge of the particular numbers and their sums and products?

I am inclined to think this is so. You are inclined to doubt it. You feel that knowledge of the numbers — which are which, and what the various sums and products are — comes both first and quite independently of knowledge of laws. I feel that knowledge of laws is implicit in our knowledge of the numbers.

I feel that laws are implicit even in an arithmetic entirely lacking the means for putting laws into words. This is connected with the importance I attach to the *systematicity* of arithmetic. But you feel that knowledge of which numbers are which, and what the sums and products are, are points of information which can be picked up one by one, in a non-systematic way, and that a system emerges only as we see certain patterns in the particular number facts we arrive at one by one.

Do I get our two perspectives right when I describe them in this way? What do you think?

P. Yes. I think so. And now it strikes me that you have perhaps moved away from your earlier view that arithmetical knowledge *could* be based on knowledge of axioms and deduction. For now you see the axioms as there in arithmetic even when the arithmetic in question cannot so much as formulate them as axioms. For what you say of laws holds for axioms, since axioms are laws.

C. Yes. It no longer is clear to me that mathematical knowledge comes down to knowledge of laws plus *deduction* from laws. The knowledge of the laws is, as it were, there from the start as knowing *how* to count, to add, and to multiply.

I would now prefer to say that talk in terms of axioms and deduction no more than *alludes* to what is going on in obtaining arithmetical knowledge — or *represents* that knowledge in a certain way. It puts things, we might say, in a «deductive style». And that might even be highly misleading presentation of that knowledge.

P. Well — I'm certainly pleased that you've finally become quite clear in your statements!

C. Yes. I too thought I was doing a lot better.

So here we are. Confused. Not a bad starting point for philosophy.

There are the *particularities* of the numbers, and the *laws*. The real issue, as I now see things, has nothing in particular to do with *axioms*. They are of interest only when one wants, as it were, to logically sum up a body of knowledge. Instead, it is the relation of the laws to the particularities that concerns us, and how law and particularity connect up with knowledge.

Or I might put it this way: We are wondering about the relation between non-quantificational and quantificational arithmetic.

So we need to ask ourselves: What are the roles of generality in arithmetic?

And here three things immediately present themselves: First, there is the generality present in our methods of calculating particularities. Second, there is the generality that enters into the formation of new mathematical concepts, and, third, there is the generality we arrive at through proof.

Charles Sayward
University of Nebraska-Lincoln
<csayward@unlserve.unl.edu>

SORITES (ΣΩΠΙΤΗΣ), ISSN 1135-1349
http://www.sorites.org
Issue #16 — December 2005. Pp. 62-73
Is Theism More Rational Than Agnosticism: A Critique of
Arguments for the Necessary Existence of God?
Copyright © by SORITES and David Kimweli

IS THEISM MORE RATIONAL THAN AGNOSTICISM: A CRITIQUE OF ARGUMENTS FOR THE NECESSARY EXISTENCE OF GOD?

David Kimweli

Introduction

In April 2003, Paul Geisert and Mynga Futrell of Sacramento California presented a strategy for atheists to accrue more political recognition, influence, and power. Their solution involved introducing a new meme¹, a term that atheists and agnostics could apply that would disassociate themselves from the negative connotations linked with the existing idiom. Many politically outspoken atheists perceived a 1999 Gallup Poll that found less than half of all Americans willing to vote for an otherwise qualified candidate for president who was an atheist as evidence of persecution. According to Michael Shermer, this Poll illustrated that «(atheists) must unite against the prejudice that as nonbelievers we are not qualified to be full participatory citizens» (2). Geisert and Futrell's answer to this image problem was to redefine non-theists, agnostics, atheists, secular humanists and the like as «Brights.» The term Bright, in this context, is «a person whose worldview is naturalistic — free of supernatural and mystical elements.» Geisert and Futrell are optimistic that the application of a pleasant new moniker to religious skeptics will de-stigmatize the community and propel them into a significant social and political force. They write, «Given our severe linguistic disadvantages, the Bright movement asks those with a naturalistic worldview to join minds and begin to view themselves and speak in civic situations as Brights» (1).

The reaction to the emergence of the Bright movement was strong. Atheists like Richard Dawkins and Daniel Dennett became enthusiastic supporters, penning columns expressing themselves as Brights in the *Guardian* of London and the *New York Times* respectively. Conversely, criticism abounded by theists and non-theists alike, largely centering on the selection of a loaded term as «bright,» of which the antonym is dim, as blanket description of all non-believers. The criticism argues, and quite convincingly, the Bright movement is counterproductive because it seemingly equates non-belief with intelligence and belief with dimness, only furthering a negative perception that atheists are arrogant and snobbish, which not coincidentally, is exactly how Dennett came across in his *NY Times* piece. In his column Dennett described atheism as intellectually superior to theism as a matter of fact. The «naturalist» worldview is more rational than its super-naturalist counterpart, Dennett claims,

¹ This term was first coined by Richard Dawkins in *Selfish gene* (1976). In this work he floated the idea that genes in the organic world might be mirrored by memes in the world of ideas, and that the spread of ideas might be understood in terms of replication, variation and differential survival (Radcliffe-Richards 53).

because it refuses to fall for silly superstitions like «ghosts, elves, or the Easter Bunny — or God» (1).

Dinesh D'sousza took great umbrage with the Dennett's previous claim, firing back in his *Wall Street Journal* editorial «Not So Bright.» In his piece, D'sousza invokes the Kantian limits on knowledge as incapable of revealing reality in its entirety. Thus, reason's inability to understand the totality of reality «opens the door to faith» or offers a sensible rationale for believing in God. He writes:

(Brights) should refrain from the ignorant boast that atheism operates on a higher intellectual plane than theism. Rather, as Kant showed, reason must know its limits to be truly reasonable. The atheist foolishly presumes that reason is in principle capable of figuring out all that there is, while the theist at least knows that there is a reality greater than, and beyond, that which our senses and our minds can ever apprehend.

D'sousza's application of Kant to advocate his theism is clever and effective for the purpose of communicating to a broad readership not well versed in the highly abstract features of Kant's ethics, but not the primary argument for God the great philosopher employed himself. Instead, Kant makes what is widely known as the moral argument for the existence of God. The moral argument reasons that while theoretical proofs of God's existence are impossible, God is necessary to preserve the attainment of a good will — or the fundamental object of a moral life. Lewis White Beck, in the introduction to his translation of the *Critique of Practical Reason* writes, «God appears as necessary to the existence of the *Summum bonum*. God is the being that guarantees happiness in proportion to virtue; and moral laws, in whose fulfillment lives man's worthiness to be happy, can be looked upon as divine commands» (48).

The debate between Dennett and D'sousza gives rise to several questions of philosophical import. Specifically, is belief in God mere superstition as the Brights claim? Also, considering Kant himself believed in the impossibility of proving God's existence, is the moral argument the best rational and intellectual justification for belief? I contend Kant's transcendental morality supplies a less adequate justification for belief than John Dewey's moral instrumentalism. Lastly, while I concede belief in God is morally instrumental, I will argue that agnosticism is preferable to theism.

Section I

Is the Existence of God Provable?

There are three traditional arguments that purport to prove the existence of God. These are the ontological proof, cosmological proof, and the physico-theological proof. Kant, while a theist, systematically discredited these three proofs in his *Critique of Pure Reason*. His counter-arguments were so effective they were considered definitive until the 20th century when some scholars attempted to resurrect the ontological proof, the most notable of which was put forth by Alvin Plantinga. Plantinga's reply to Kant will follow shortly, but first let us summarize Kant's criticism of the ontological proof.

The ontological proof has taken several forms over the centuries. St. Anselm first made the particular version of the proof to which Kant objected. Anselm contended that the ability to conceive of a greatest possible being, or God, makes His existence necessary as the concept of «greatest» automatically entails existence. In other words, to insist the greatest possible being is only a concept brings forth a contradiction since if God actually existed, He would be greater than the conception.

Kant's objections with Anselm's version of the ontological proof are with its logic. According to Kant, no contradiction arises in denying the subject and predicate of a synthetic existential proposition. In the case of analytic propositions, or those in which the content of the subject is assumed in the predicate, like «this triangle's angles amount to 180 degrees,» no contradiction arises from denying the existence of the triangle as well as its three angles. If, on the other hand, one would accept the subject while denying the predicate, as in maintaining, «this particular triangle did not have three angles,» a contradiction occurs. However, the claim «There is no God» does not involve a contradiction since the subject God and its existential predicate are simultaneously rejected (*Critique of Pure Reason* 502).

Kant makes clear the impossibility of the ontological proof when he writes the following, «Whatever, therefore, and however much our concept of an object may contain, we must go outside it if we are to ascribe existence to the object... (the concept) is altogether incapable, by itself of enlarging our knowledge in regards to what exists» (504) Even if Anselm's conception of a greatest possible being implied existence by definition, the matter of whether this being exists or not is only verifiable in a context apart from the assertion of the proposition. Simply put, Kant insists the ontological proof mistakes internal validity as evidence of a proposition's truth or falsity. «A greatest possible being exists» is not a contradictory proposition. Nevertheless, «there is no greatest possible being» is equally non-contradictory. Truth or falsity is a further question decided empirically, and not by the assertion of existential propositions.

Kant's criticism of the impossibility of the ontological proof was considered definitive until the recasting of the proof by contemporary scholar Alvin Plantinga. Plantinga employs modal logic in an attempt to make the ontological proof, at the very least, rationally acceptable. Plantinga's version of the proof invokes world-indexed properties to reformulate the Anselmian claim that existence in reality is greater than existence in thought alone, while skirting the Kantian criticism that the ontological proof fails because it predicates existence (Mackie 55). Plantinga cites the possibility of objects existing in logically possible «other worlds» apart from the actual world as illustrating an object of maximal greatness must exist. While Plantinga's proof is sprawling, the gist of his argument is reducible to the following 4 premises.

1. It is possible that there is a greatest possible being.
2. Therefore, there is a being that in some logically possible world has a maximum degree of greatness — a degree of greatness nowhere exceeded.
3. A being *B* has the maximum degree of greatness in a given possible world only if *B* exists in every possible world.
4. A maximally great being *B* exists necessarily in the actual world if it is possible in every possible world (105-106).

Plantinga makes a crucial leap in his reasoning from (3) to (4), and it is that logical possibilities are identical for every possible world. He writes, «If a given proposition or state of affairs is impossible in at least one possible world, then it is impossible in every possible world» (106). I think the previous is an article of faith that Plantinga masquerades as a logical rule. A quick glance at the concept of «logically possible other worlds» casts doubt that impossibilities hold across all possible worlds. According to Patrick Hopkins, «logically possible other world» simply means «Is there a set of conceivable conditions so that *X* could

have been the case instead of *Y*? If so, *X* is logically possible.» If the previous is what the concept of logically impossible entails, then a paradox surfaces if one holds that what is impossible in one logically possible world is impossible for all worlds. The paradox is the logical impossibility that the earth never existed in the actual world. It is given that there is an earth in the actual world, and the actual world is included in the set of all logically possible worlds. But despite the impossibility of earth never existing in the actual world, it is conceivable that the volatile conditions under which the universe took form could have produced a universe in which the earth never existed. It is entirely logically possible that the cosmos could have never brought forth our planet, yet the non existence of earth is impossible in the actual world. The previous argument is better explicated in the following form:

1. The actual world exists.
2. The actual world is logically possible.
3. It is logically impossible that this world never existed in the actual world.
4. There are logically possible other worlds where this world never existed.
5. Thus, what is logically impossible in the actual world does not necessarily hold true for all other worlds.

The preceding proof is a *reductio ad absurdum* of the Plantinga proposition «If a given proposition or state of affairs is impossible in at least one possible world, then it is impossible in every possible world.» What the proof demonstrates is the impossibility of anything not being itself. This impossibility, however, can only apply to the actual world since possible other worlds and all possible entities within them by definition are hypothetical and nonexistent. So there are instances where what is impossible in the actual world is logically possible in other worlds, which illustrates step (4) in Plantinga's proof does not follow from (1-3), and thus is, as a consequence, an insufficient attempt to make a rationally permissible ontological proof.

Kant's criticism of the cosmological proof or first cause argument is similar to his objection with the ontological proof, mainly because he thinks the former presupposes the latter. Kant writes that the cosmological proof «retains the connection of absolute necessity with the highest reality, but instead of reasoning, like the (ontological) proof, from the highest reality to necessity of existence, it reasons from the previously unconditioned necessity of any being to the unlimited reality of that being» (508). The cosmological proof reasons that since contingent being exists, they must have been caused by necessary first cause. For one, the concept of a necessary cause does not imply that it is maximally great, or possesses any of the properties theist commonly attribute to God. As A.G. Ewing explains in his *Short Commentary on the Critique of Pure Reason* that «to establish the kind of God theologians want, the ontological argument is presupposed» (245). Another objection Kant makes is that the cosmological proof supposedly reasons from experience the necessity of a first cause, yet «experience is unable to demonstrate this necessity as belonging to any determinate thing» (*Critique of Pure Reason* 510). The only way, then, the cosmological proof gets God out of its argument is to make the assumptions of the ontological proof, and subsequently all the criticisms of the ontological proof apply to the cosmological one.

Finally, Kant explains the physico-theological proof the following way: «If, then, neither the concept of things in general nor the experience of any existence in general can supply what is required, it remains only to try whether a determinate experience, the experience of

things in the present world, may help us to attain an assured conviction of a supreme being» (518). The physico-theological proof reasons that the order and purpose observable in the world infers intelligent design, like William Paley made in his famous watch argument. Paley writes, «Every indication of contrivance, every manifestation of design, which existed in the watch, exists in the works of nature — the inference, we think, is inevitable, that the (world) must have had a maker» (Pojman 669). Kant, although conceding the physico-theological argument is the proof that is «the oldest, the clearest, and the most accordant with the common reason of mankind» it fails because «the physico-theological proof can never by itself establish the existence of a supreme being, but must always fall back upon the ontological argument to make good its deficiency» (521). As a result of the previous, the inadequacies of the ontological proof pertain to the physico-theological one.

Kant's objection to the ontological, cosmological, and physico-theological proofs for God leads us into our next question: given the inability to prove the existence of God empirically or logically, are there other rational reasons for believing in God? Kant, while skeptical of theological proofs, was a devout believer and convinced God was necessary for the goal of a moral life — the triumph of a good will.

Section II

Is God Necessary to Preserve Morality?

The argument Kant sets forth in his *Foundations of the Metaphysics of Morals* and the *Critique of Practical Reason* is essentially that morality needs justification in universalizability, reason requires its position in the «categorical imperative,» or unconditional moral law. For instance, Kant writes, «For law alone implies the concept of an unconditional and objective and hence universally valid necessity — the categorical imperative is restricted by no condition. As absolutely, though practically, necessary it can be called a command in the strict sense» (*FMM* 76).

The need for universality in moral theory for Kant grows out of his demarcation of a theoretical² boundary of knowledge that renders it incapable of grasping ultimate reality; or knowledge apart from the categories our mind employs to synthesize input; or knowledge. Our mental apparatus is inadequate for obtaining the essence of reality because the categories frame the context by which we can know an object while objects exist as things-in-themselves. Since we can never escape the context by which our mind frames sense data on objects, or get «outside» the proverbial box, knowledge of things-in-themselves is not possible.

While acknowledging the impossibility of knowing things-in-themselves Kant recognizes that the rational curious mind will inevitably conjecture as to the nature of ultimate reality. Three ideas about ultimate reality often supposed by the rational mind are freedom, God, and immortality — each of which plays key roles in Kant's moral theory — freedom is exercised by a will, of which an irrevocably good one is the *Summum bonum* or object of a moral life. However, Kant believes an irrefutably good will is impossible to attain in practical experience, and thus immortality and God, as objects of pure reason, are needed to achieve it. Kant explains as much when he writes:

² Kant's Critiques are famous for the distinction between theoretical and practical knowledge. Theoretical knowledge encompasses the principles of reason and a *priori* conditions of the mind that are constitutive for all experience. Practical knowledge involves the principles of the understanding or the manner in which perceptions are synthesized into objects of experience.

The ideas of God and immortality are, on the contrary, not conditions of the moral law, but only conditions of the necessary object of a will which is determined by this law, this will being merely the practical use of our pure reason— thus, through the concept of freedom, the ideas of God and immortality gain objective reality and legitimacy and indeed subjective necessity as needs of pure reason (119).

What Kant means is that given that a will in compliance with the moral law, a good will, is the object of a moral life, God and immortality are required to make this possible. But before I explain this point further, it is important to clarify how Kant thinks we can have moral experience. Kant argues that non-mechanical causality, or freedom, is necessary for morality because it is the only way moral experience is even possible. Kant writes:

The determination of the causality of beings in the world of sense as such can never be unconditioned, and yet for every series of conditions there must be something unconditioned, and consequently a causality which is entirely self-determining. Therefore, the idea of freedom as a faculty of absolute spontaneity was — as far as its possibility was caused, an analytical principle of pure speculation — thus the concept of freedom is made the regulative principle of reason (158).

The antinomy of causality, the conflict of whether the causation is better explained in terms of non-mechanical or natural, then, is settled by applying mechanical natural causality to objects-in-themselves, and non-mechanical free causality to actions by «objects of understanding,» or rational beings. According to Lewis White Beck, this dichotomy is based on Kant's unwavering belief in the necessary and universal maxims of geometry and natural physics. He states, «Kant's conviction of the certainty of Newtonian Mechanics was too deep to be shaken by any negative conclusion drawn from speculations concerning the human mind... (assured he was) of the validity and certainty of the synthetic judgments of geometry and physics» (10). Of course, special relativity theory illustrates the limitation of Newtonian physics to predict laws of motion only on the unique frame of reference found on earth, while inadequate to predict physical behavior when applied to motion at speeds approaching light. The non-universality of Newtonian Mechanics in turn casts doubt on Kant's belief in mechanical causality, but that is not the issue that concerns us at the moment. For the time being, let us restrict ourselves to the implications freedom holds for moral action. As Kant demonstrated, freedom is necessary for the possibility of moral experience, or else human behavior is the determined outcome of a mechanical process.

Even more than the basis for moral experience, the idea of freedom is rationally permissible when natural causality is confined to phenomena and non-mechanical causality (freedom) to super-sensuous reality. H.W. Cassirer in *A Commentary on Kant's Critique of Judgment* explains Kant's resolution of the antinomy of causality the following way, «If absolute reality be ascribed to the world of appearances, freedom cannot be upheld. It is only if we allow that there may be things-in-themselves which are independent of our world of sense and the laws that determine it that we can reasonably assume the existence of a supersensible principle of freedom» (58). In spite of the fact that this distinction between natural causality and objects of experience from freedom and objects of the super-sensuous prevents a contradiction, Kant asserts that theoretical knowledge of freedom or intelligible causality is impossible. Cassirer summarizes Kant's position when he writes:

The only kind of causality of which we have any knowledge is natural causality, which is a rule of the understanding for the determination of natural events. We cannot understand how there can be a causality of an entirely different kind; nor do we need to understand this, for it does not concern us as moral agents since we cannot and need not have any knowledge of the supersensible moral law (64).

Kant's conception of the moral law is also divided into theoretical and practical knowledge. As stated in the citation above, Kant does not think theoretical knowledge of the

moral law is possible or even relevant. The moral law is revealed practically, however, in experience. To explain how we can develop maxims, or rules to discern the consistency of a particular act with the moral law, Kant contends the action must be universalizable, respect human dignity and be self-revealing through reason.

In his *Foundations of the Metaphysics of Morals*, Kant states:

And what is it that justifies the morally good disposition or virtue in making such lofty claims? It is nothing less than the participation it affords the rational being in giving universal laws. He is thus fitted to be a member in a possible realm of ends to which his own nature already destined him. For, as an end in himself, he is destined to be legislative in the realms of ends, free from all laws of nature and obedient to these which he himself gives. Accordingly, his maxims can belong to a universal legislation to which he is at the same time also subject (93).

Finally Kant believes that a good will, although only the result of duty and obligation to the moral law is nevertheless conducive to happiness, a happiness which is rewarded proportionally to virtue. To put it slightly differently, the notion of the highest possible good entails happiness. Yet, since natural causality has no explanatory power as to the motives of the will, since the will itself is an example of non-mechanical causality, the absolutely good will and thus the highest possible good must have a supreme cause not found in nature, chiefly God. Kant puts forth this moral argument for God when he writes:

Therefore, the highest good is possible in the world only on the supposition of a supreme cause of nature which has a causality corresponding to the moral intention. Now a being which is capable of actions by the idea of laws is an intelligence, and the causality of such a being according to this idea of laws is his will. Therefore, the supreme cause of nature, in so far as it must be presupposed for the highest good, is a being which is the cause of nature through understanding and will, i.e., God — therefore it is morally necessary to assume the existence of God. (*Critique* 228).

In summary, Kant's transcendental moral theory needs to God to preserve the possibility of the highest good, which is proportional happiness ensuing from compliance with the moral law. The problem with this position is, as I see it, its preoccupation with the notion of an indubitable moral law. Kant thinks that the moral law is an indisputable fact. In the *Critique of Pure Reason* he writes, «The moral law is given as a fact of pure Reason of which we are a priori conscious, and which is apodictically certain» (522). The concept of a moral law hinges on the view morality was configured as opposed to developed. I contend that an instrumentalist conception of morality, one that views morals as practical developments that arose in the progress of human affairs, is not only a simpler explanation but one inexorably more suited to the adaptive needs of the human condition.

Section III

Is an Instrumentalist Moral Theory Superior to a Transcendental One?

While Kant's critical system departs from the rationalist and empiricist schools by arguing against human capacity to attain knowledge of ultimate reality, he laments this impossibility and addresses ways to cope with it. In other words, knowledge of ultimate reality does not lose its place as the most central and prime aspect of knowledge under moral transcendentalism, albeit an impossible one.

John Dewey in his *Reconstruction in Philosophy* takes an entirely different approach to knowledge. Ralph Ross, in the introduction to his reissuing of the *Reconstruction* writes of Dewey's outlook on knowledge, «It is not a fixed form, essence, or structure behind processes of change that is the key to knowledge, but the way change becomes evident as we experiment

with things.» He goes on to write, «Knowledge is not a matter of discovering what things ‘really are’, as though they were unchanging and inhabited a universe without us, they are what they can do and what can be done with them» (xix). This is a synopsis of what is known as an instrumentalist theory of knowledge — what is paramount to *know* is what we can use to transform and shape experience for our benefit; as opposed to the Kantian transcendentalist viewpoint that places undiscoverable characteristics of objects (those aspects divorced from the interpretative activity of the mind) as ultimate. The Deweyan perspective is that moral knowledge is like any other area of human inquiry, it useful to the extent it has adaptive and transformative power in shaping human experience. Richard Rorty summarizes Dewey’s moral instrumentalism in the following way, «Physics is a way of trying to cope with various bits of the universe; ethics is a matter of trying to cope with other bits — the question of what propositions to assert, which pictures to look at, what narratives to listen to and comment on and retell, are all questions about what will help us get what we want» (xliii). «What we want» from morality is a set of guidelines to human affairs that will bring about the most tolerable living conditions; and in the course of history, civil, fair, and respectful interaction between persons in a society and assorted societies have been most favorable to bringing forth the most unobjectionable habitation. In this sense morality is grounded in experience; however it is not subject to the criticisms of limitation Kant makes. The moral maxims derived through instrumentalism are valid and applicable because they are proven to have brought about a suitable environment. Since the human condition is unstable and changing instrumentalist moral maxims have gone under alteration, but these adaptive measures were not corrupting to the extent they were undertaken to the end of improving the state of affairs by achieving civil, fair, and courteous treatment of fellow persons and among cultures.

Dewey describes instrumentalism in this manner, «The interaction of organism and environment, resulting in some adaptation which secures utilization of the latter, is the primary fact, the basic category — knowledge is not something separate and self-sufficing, but is involved in the process by which life is sustained and evolved» (79). The instrumentalist objection Dewey raises illustrates that there is really nothing to know about things-in-themselves, since they are incomprehensible and offer nothing in the way of aiding us adapt, transform, and improve our environment. As a result, knowledge should be applied to aspects of experience we can use to further develop and progress the human condition. In short, ‘science, as Dewey understood it, was revolutionized by the discovery that what was «universal» was process (xix).

Instrumentalism has vastly distinct import for morality than transcendental moral theory. Whereas Kant sought to ground morality in unchangeable and fixed moral truths that underlie any good action, Dewey saw morality as the application of intelligence or scientific methodology to practical human affairs. The consequence of Dewey’s theory, then, is not an unalterable law but regulations of conduct variable to the adaptive pressures that will inevitably surface over time. Dewey sets forth his most effective criticism of Kant’s transcendental moral theory when he states:

After all, then, we are only pleading for the adoption in moral reflection of the logic that has been proved to make for security, stringency and fertility in passing judgments upon physical phenomena. And the reason is the same. The old method in spite of its nominal and esthetic worship of reason discouraged reason, because it hindered the operation of scrupulous and unremitting inquiry (174).

The moral knowledge obtained in an instrumentalist theory, therefore, has utility and flexibility, which I contend is more suited to the needs of the human condition, a condition which is ever transforming.

Section IV

Absent the Moral Argument, why Believe in God?

As we have seen, God props up Kant's moral argument, as He is the supreme cause of «the highest good.» In contrast, while Dewey's moral instrumentalism does not do away with God; it renders him/her unnecessary. God is not necessary in an instrumentalist moral theory because humans developed morality to provide the most acceptable means of co-existing with one another. While some have been less developed or «instrumental» than others, every civilization throughout history has established a code of conduct, or rules that dictate standards of decency in the course human affairs.

Nonetheless, moral instrumentalism does not eliminate God. In fact, the existence of God is not an important question under moral instrumentalism. What matters under the instrumentalist view is whether or not belief in God is conducive to bringing civility in personal and communal interactions. Historically, to say the least, believers have had mixed results in behaving civilly towards non-likeminded groups.

Yet, the strife borne from conflicts between believers, at least as observed among the followers of the Abrahamic deity (the one to which most of humankind proscribes), is arguably more a corruption of faith so far as its proper manifestation encourages basic dignity among persons. It seems that belief in God, then, is worthwhile in so far as it motivates its espousers to live morally. In the past, belief in God has achieved the previously stated goal when it conjures conviction without dogma. Indeed, it is a strident dogma among fanatics in their particular idiosyncratic and sectarian God that induces intolerance. However, the immoral acts of fanatics in the name of religion do not detract from the charity and good works of devout believers who have furthered the aim of a moral society.

Belief in God has worked at times in fostering moral values in humanity, and, as a result, is worthy of respect. The condescension of the Brights to all belief in God is unwarranted for the reason that they fail to take into account all the good and decency that was and is still motivated by sincere faith. While it is true dogmatic believers are responsible for much intolerance and immorality, such as the 9/11 hijackers, it is undeniable that belief has brought about many positive outcomes as well, and I will venture to say, more good than harm.

It is of relevance to reveal that while I find instrumentalist use for belief in God, I am not a theist, but an agnostic. The version of agnosticism to which I subscribe divides all affirmations of opinion into informed and uninformed. An informed opinion is one predicated on and corroborated by facts. In contrast, uninformed opinions are inconsistent with or unverifiable pertaining to the facts. Before I further explain whether belief in God is informed or uninformed, I want to make clearer what actually constitutes a fact. A fact is a statement that conforms to the grammatical precepts of the language it is expressed and whose vocabulary research, investigation, and inquiry dictate as warranted to assert.

The intense study of language Psychology «turned» to in the 20th century, made breakthroughs with an approach to language as behavior governed by rules. The language as behavior position was first outlined by Norm Chomsky in his landmark *Syntactic Structures*. Chomsky's theory defines language as rule conforming behavior, or as Neil Smith and Deirdre

Wilson write, «language is best described in terms of a grammar, or system of rules» (*Modern Linguistics: The Results of Chomsky's Revolution* 21). Whereas there are slight variations in the rules, as evidenced in the alternate pronunciations and meanings of words and phrases across dialects, the observation of fundamental grammar is necessary for any understandable linguistic expression. The psychological reality of grammar principles as necessary for the possibility of linguistic comprehension is the most logical explanation for the ability of language users to «form judgments about utterances that they have never heard before» (22). To illustrate that humans indeed have this capability, I offer this sentence, «The Cheetah gazes in the river valley of Machakos, Kenya.» It is unlikely that you have ever heard the previous string of words spoken precisely in the succession they appeared, yet your unfamiliarity with that sentence did not prevent you from assessing its grammatical correctness. As a consequence then, the determination of the grammatical or ungrammatical status of novel sentences is best explained as the application of a «grammar which embodies the principles of sentence formation and interpretation.» Therefore, for a statement of fact to be understandable, one criterion is that it must conform to the rules of sentence structure and formation in the language it is expressed.

The second criterion that any statement of fact must meet is that other users of a given language will find the particular assertion warranted, or justified. The term justified in this limited sense only means that most users would concede or agree to the point of information offered as a fact, and does not contend the statement of fact «refer» to objects in the world or «represent» them in a positivist way. This definition restricts the conferring of fact status to a vocabulary of description most users in a language agree is justified. Agreement on which propositions are indeed facts is developed through a laborious process of investigation and inquiry, and ultimately gauged by the explanatory power of the descriptive vocabulary. To explain more fully, a vocabulary begins as an attempt to describe or explain some aspect of our environment. If repeated inquiry confirms the hypothesis, the users of the language reach a consensus that the statement is warranted to assert, and is thus a fact.

So, a statement of fact is both grammatically consistent and vindicated through inquiry. Francis Bacon was the first to set forth that humans induce generalizations, or statements of opinion, which are deemed warranted or unwarranted to the extent they follow from the pertinent set of facts. The logical extension of the inductive method is that an opinion is informed if it is consistent with the facts, and uninformed if contradictory or unverifiable to the facts. That an opinion contrary to the facts would be uninformed is evident enough, but the unverifiability criterion deserves more clarification. If the conference of fact status on a vocabulary is determined by the degree it is warranted, then only statements that are confirmed through inquiry could ever attain fact standing. Thus an unverifiable proposition, incapable of validation through inquiry, can never be a fact.

My agnostic argument takes issue with how opinion on the existence of God is typically settled. Theists often assert that «God exists» is warranted because of revelation, logical necessity, or moral necessity. I have spent a good deal of time in this paper arguing against the latter two arguments, but I am even more skeptical of revelatory claims for God than the logical or moral ones. Such claims, as explained by David Hume, are entirely dependent on testimony and incapable of any corroboration with evidence (Pojman 710). The claim «God spoke to me» is one on which the testimony of the speaker alone is the only means to validate such communication actually occurred, and thus not achieving the status of factual evidence, a position conferred by investigation.

Since the statement «God exists» does not have any factual basis, it is an uninformed opinion. It should be noted that I am applying the terms ‘informed’ and ‘uninformed’ in an extremely limited sense, and am in no way denying the moral instrumentality of belief in God. An uninformed opinion strictly means one that contradicts the facts or is unverifiable with facts at this time. The uninformed categorization does not eliminate the possibility that at some future point evidence could surface which verifies the existence of God, it simply contends no such evidence is available presently, and so contemporary existential claims for God are not warranted on a factual basis. My unbelief results from my hesitation to assert opinions I cannot support with facts. I do not share Dennett and the Brights disdain for theism because I recognize its moral instrumentality, but consider it a less favorable position to agnosticism because it makes a fact-less claim. The agnosticism I hold simply refuses to make a belief claim about God, roughly stating, «The existence of God is a question currently beyond verification, and consequently an unwarranted assertion.» Therefore, since agnosticism is not inconsistent with any known facts, it is preferable to the theistic alternative, which makes an existential claim in the absence of factual evidence.

Works Cited

- Cassirer, H.W. *A Commentary on Kant's Critique of Judgment*. New York, Barnes & Noble, Inc., 1970.
- D'souza, Dinesh. «Not So 'Bright'.» *Wall Street Journal* 12, October 2003: A 16.
- Dennett, Daniel. «The Bright Stuff.» *New York Times* 12, July 2003: A: 19.
- Dewey, John. *Reconstruction in Philosophy*. Ed. Jo Ann Boydston. Carbondale and Edwardsville: Southern Illinois University Press, 1988.
- Ewing, Daniel. *A Short Commentary on Kant's Critique of Pure Reason*. Chicago: University of Chicago Press, 1938.
- Kant, Immanuel. *Critique of Practical Reason: And Other Writings in Moral Philosophy*. Trans. Lewis White Beck. Chicago: University of Chicago Press, 1949.
- Kant, Immanuel. *Critique of Pure Reason*. Trans. Norman Kemp Smith. New York: St. Martin's Press 1965.
- Mackie, J.L. *The Miracle of Theism: Arguments For and Against the Existence of God*. Oxford: Clarendon Press, 1982.
- Plantinga, C. Alvin. *God, Freedom, and Evil*. Grand Rapids: William. B. Eerdmans Publishing Co., 1977.
- Pojman, Louis P. *Classics of Philosophy*. New York: Oxford University Press, 1998.
- Richards, Radcliffe Janet. *Human Nature After Darwin*. Routledge: New York, 2000.
- Rorty, Richard. *Consequences of Pragmatism*. Minneapolis: University of Minnesota Press, 1982.
- Shermer, Michael. «The Big 'Bright' Brouhaha: An Empirical Study on an Emerging Skeptical Movement.» <<http://www.the-brights.net>>

Smith, Neil and Deirdre Wilson. *Modern Linguistics: The Results of Chomsky's Revolution*.
Bloomington: Indiana University Press, 1979.

David Kimweli
<weli37@hotmail.com>

RULES AND REALISM: REMARKS ON THE POVERTY OF BRUTE FACTS

J. Jeremy Wisniewski

I. Brute and Institutional Facts

In 1966, Peter Berger and Thomas Luckmann published a book that to this day constitutes a definitive text in the sociology of knowledge: *The Social Construction of Reality*. By developing the implications of the phenomenology of Alfred Schutz, as well as the philosophical anthropology of Helmuth Plessner and Arnold Gehlen, Berger and Luckmann effectively provided a theoretical perspective through which we might examine the emergence, maintenance, and manipulation of discrete knowledge-systems as they emerge in the symbolic universe of culture.

Nearly 30 years later, in *The Construction of Social Reality*, John Searle published a sustained riposte. In it, he claims that it is possible to distinguish brute and institutional facts. Institutional facts are those facts that are facts only through collective intentionality. According to Searle, these facts are only possible in light of 'brute' facts—i.e. in light of some reality independent of all human agency. The appropriate analysis of institutional reality—of the socially constructed world—reveals that there are things which cannot be so constructed; that there are things which *must* exist independently of human intentionality. In what follows, I argue that Searle conflates two separate theses in his defense of External Realism, and that this leads him to beg the question against the anti-realist. One can acknowledge Searle's ER while also holding that Searle inverts the relation of dependence between brute and institutional facts. To put the response succinctly: we can call some fact 'brute' only because of entrenched institutional facts. The very existence of 'brute facts' can itself be regarded as an institutional artifact.

II. The Argumentative Leap: Searle on External Realism

Searle claims that institutional facts have the logical form of constitutive rules ('x counts as y in context c'). The existence of money, to take one example, is analyzable in terms of a rule, or set of rules, which describe what money *is*. Thus, we might say that 'this piece of green paper (x) counts as money (y) when it is made in the appropriate way, by the appropriate agency, and so on (in c).' As is obvious, this analysis itself contains further institutional facts: 'a piece of paper' is not part of the furniture of the universe. Rather, it is the product of human needs and purposes. Thus, as Searle rightly acknowledges, a complete analysis of any institutional fact via constitutive rules is likely to yield iterations of rules: 'this material counts as a piece of paper when it is used in thus and such a way' could thus take the place of the original x term, making the constitutive rule have the following structure: (x

counts as *y* in *c*) counts as *y* in *c*.¹ At the end of such iterations, Searle contends, institutional facts bottom out in brute reality: there is some brute '*x*' upon which we impose some status '*y*.' Without an *x* upon which to impose some status function, there would be nothing out of which social facts might be constructed.

One problem with this view (and there are many) is that it begs the questions against the anti-realist. Searle has simply helped himself to a realism about brute facts in order to establish his account of social ontology.² Without establishing that there are objects independent of our systems of descriptions—and hence brute—it is open to the anti-foundationalist to contend that there simply *are no* brute facts.

At the end of *Construction*, Searle grants that it is time to pay his argumentative dues: he claims that he will provide an argument for a view he calls 'External Realism' (hereafter, 'ER'). The argument given is transcendental: it aims to articulate the presuppositions of some *x* which has been taken for granted (in this case, language-use). There are two things I want to claim regarding Searle's argument: 1) it works, and 2) it is irrelevant to determining whether or not there are brute facts. To do this, it will be useful to spend a moment considering the nature of transcendental arguments. This will allow us to see what the status of Searle's external realism is, and also why it is irrelevant to the discussion of brute facts.

There are several more or less loose ways of articulating the significance of the results of transcendental arguments. We can call the results of such arguments constitutive rules of experience, or discoverable analytic statements, or necessary and sufficient conditions of experience, or propositionally expressed pieces of a world-picture, or, finally, with the old wisdom, the conditions for the possibility of some *x*. All of these are but casting shadows: rather vague ways of spelling out that at which a transcendental argument aims — which is, to state it oversimply, what *must* be taken for granted in the employment of a concept.

This expression of transcendental arguments, however, begs for a prepositional phrase: must be taken for granted *for what purpose?* To state the answer boldly: transcendental arguments yield intelligibility conditions; they express what makes certain concepts we possess intelligible. A transcendental argument tells us what we must take for granted if some *x* is to remain intelligible. Because transcendental arguments begin with a concept already in our conceptual panoply, the result of a transcendental argument cannot be epistemic gain: we do not learn anything about the furniture of the universe from intelligibility conditions; rather, we clarify our conceptual situation.

To offer a quick example: the solipsist who articulates her philosophical view is subject to transcendental refutation in the following way. One first shows that the meanings of particular terms in a language are public. The notion of each word having a meaning for each individual singularly is, the argument would go, preposterous. This means that, in order for the solipsist to express her view, she has to rely on the meanings of the terms she uses. Because she relies on these meanings, she has already conceded that there is a 'public' — namely, other language-users with whom she might speak. But this means precisely that, in

¹ It would also be necessary to provide some such rule for green, as well as any other social facts which appeared in the iterated rule.

² Moreover, it isn't even clear that the analysis Searle offers *requires* realism about facts. This will become clear in what follows.

order to say that I am the only thing that exists, I have to presuppose that I am not the only thing that exists. Solipsism turns out to be a performative contradiction on this view. The articulation of this position ignores a condition for the possibility of meaningful assertion.³

At the end of *The Construction of Social Reality*, John Searle offers an argument against a caricature of anti-foundationalists.⁴ He (rightly) claims that the argument he invokes is a transcendental one. It runs as follows:

1. The normal understanding of utterances in a public language requires that the utterances be understandable in *the same way* by any competent speaker and hearer of the language.
2. A large class of utterances purport to make reference to phenomena that exist outside of, and independently of, the speaker, the hearer, and their representations, and indeed, in some cases, independently of all representations.
3. Features 1 and 2 require that we understand the utterances of many of these sentences as having truth conditions that are independent of our representations. By purporting to make reference to *public phenomena*, phenomena that are ontologically and not merely epistemically objective, we presuppose that the truth or falsity of the statements is fixed by how the world is, independently of how we represent it.
4. But that presupposition amounts to the claim that there is a way things are that is independent of our representations, and that claim is just (one version of) external realism (CSR, 188).

The point of this argument is to show that even the phenomenal idealist and the social constructionist, to use Searle's terminology, must accept some (rather weak) version of external realism: namely, the view that there is something over and above our systems of categorization which *gets classified*. The method of the argument is straightforwardly transcendental: «We assume that a certain condition holds, and then try to show the presuppositions of that condition,» (CSR, 183). The condition which Searle takes for granted, and for which he owes thanks to Wittgenstein, is that successful communication does occur: whatever the problems of realism, we communicate with one another, and we make claims with truth values about things which are independent of us — which are in the world — whatever the status of these claims turns out to be. On Searle's view, in order for such a practice to be intelligible — in order for communication to occur when we describe objects (for example) in the world around us — we must take for granted that there is a way that the world is in virtue of which what we say can be either true or false.

But the crucial point here is not about the claims we make, but about the practices we share when making such claims. To understand what you are saying when you say, e.g., 'My dog has fleas,' I must presuppose that your words are intelligible to me — that they are not merely private representations. If I could not be sure that you were talking about a thing (a dog) which existed in the world, and which I understood in the same way (or more or less the

³ A silent solipsist, one might say, is the best kind. The above argument does not show that solipsism is wrong. It only shows that one cannot coherently express this position.

⁴ It is a caricature precisely because any anti-foundationalist could accept the (minimal) claims Searle defends at length in that text.

same way) as you do, I could not understand your utterance. But we do understand each other, and this suggests that we are presupposing that there is a world out there to which certain utterances refer, and that this world is independent of our private representations of the world:

Normal understanding requires sameness of understanding by both speaker and hearer, and sameness of understanding in these cases requires that utterances of the referring expressions purport to make reference to publicly accessible reality, to a reality that is ontologically objective. But the condition on public accessibility to the sorts of phenomena in these examples [e.g. 'My dog has fleas,' 'Hydrogen atoms each contain one electron,' 'Mt. Everest has snow and ice near the summit'] is that the way that things are does not depend on your or my representations (CRS, 186).

I find this argument compelling, but do not think anything Searle says is incompatible with, say, Foucault's philosophical approach—or with Berger's sociology of knowledge, for that matter. What I am interested in is the status of the argument Searle has given. He claims to have shown that all of us, insofar as we communicate meaningfully with one another, must take some version of external realism for granted (namely, that the world exists independently of our representations). To put the point in the idiom of the old wisdom, external realism is a condition for the possibility of (much of) our language. The existence of a world that is independent of our representations of the world is that which makes intelligible a large class of utterances we make, as well as the fact that we understand one another in the normal course of things. But what has the argument shown? Seeing the limits of the scope of transcendental arguments is actually a virtue of Searle's account:

There is nothing epistemic about the arguments. I am not saying that in order to know the truth of our claims we have to presuppose realism. My argument is completely independent of questions of knowledge or even of truth. On my account, falsehood stands as much in need of the real world as does truth. The claim, to repeat, is about conditions of intelligibility, not about conditions of knowledge (CSR, 195).

So it is that transcendental arguments do not constitute epistemic gain, on Searle's view. Searle has not shown that the solipsist is wrong about the nature of reality. In fact, he has shown nothing *whatsoever* about reality, and hence the argument has not improved our epistemic position via revealing the truth of some set of statements. The point of Searle's argument, and of transcendental arguments generally, is to clarify what is involved in some feature of human experience — in this case, the meaningfulness of particular sorts of language-use. The irony of Searle's conclusion is that the 'ism' attached to 'real' is no longer an ism at all: «External Realism is thus not a thesis nor an hypothesis but the condition of having certain sorts of theses or hypotheses,» (CSR, 178).

Our question thus becomes: does the transcendental argument used to show that external realism is a necessary postulate also support the existence of brute facts? The answer, it seems to me, is necessarily negative. The claims made for ER are non-referential—that is, one can advocate Searle's ER without thinking that any particular claim actually refers to any particular object (or group of objects) in the world. We necessarily *take for granted* that our language refers to some observer-independent reality, but we are in no position to say of any particular utterance that it does *in fact* refer to this *Welt an sich*. To infer from the fact that our entire language must presuppose reference that a particular utterance refers to the world is simply fallacious. Indeed, this is the basic error we train undergraduates to recognize under the heading of 'the fallacy of division'—attributing a predicate that is true of a set to particular members within that set.

To put this point another way: the claim that language is intrinsically referential is compatible with the claim that there are no brute facts. For this reason, one cannot infer the

existence of such facts from the referential role of our language. The non-existence of brute facts and an intrinsically referential language are compatible, quite simply, because language is fundamentally a social enterprise. We live in a world populated by social objects. It is crucial that we are able to refer to these objects—indeed, if we could not, socialization would be impossible. But it does not follow from the fact that we refer to things that we take to be unchangeable that those things *are* in fact unchangeable. The fact that we *regard* some facts as brute does not make them so.

III. The Necessity of the Institutional

I now want to establish, in a bit more detail, why Searle's argument for ER cannot establish any of the brute facts he postulates. I will confess, at the outset, that the failure of Searle's argument seems to me so obvious that I have some hesitation in articulating it—the same hesitation one has whenever one is put in a position to articulate the obvious. Nevertheless, while there has been much criticism of the apparatus Searle employs in *Construction*, there has been little comment on the argument for ER, and why it fails to justify brute facts.⁵

Having said that, I want to make clear that the failure of the argument is *not* a failure to show anything whatsoever. Searle's argument, as I boldly claimed above, *works*. The problem is that what Searle establishes as 'external realism' has nothing *at all* to do with the brute facts he employs throughout *Construction*.

Searle claims that institutional facts depend on brute facts. There are two different theses here. The first—and the one Searle utilizes in the beginning of *Construction*—is the view that there are facts that are language-independent; that there are propositions, representing states-of-affairs that would still exist even if the proposition expressing them did not. As examples, we get «My dog has fleas,» «There is snow on top of Mt. Everest,» and so on. The second thesis—and the one which we find at the end of *Construction*—is that the formulation 'x counts as y in context c' depends on their being an x upon which agents can impose a status function. In brief, institutional facts require a world of non-institutional items that *get understood* in a particular manner. This, in turn, is a condition for the possibility of language-use.

By conflating these two theses, Searle effectively begs the question against the anti-realist. The second version of ER is compatible with any anti-realist position, as it makes absolutely no substantive claims about the external world. Indeed, it seems simply to be a reproduction of the (Kantian) view that we necessarily presuppose a world independent of our representations, even if we can say nothing at all about this world. The problem with Searle, of course, is that *throughout Construction* he never tires of making claims about the content of the 'real' world (invoking the first thesis articulated above). For the sake of clarity, let's distinguish the two theses as follows:

ER₁ The world consists of states-of-affairs, accurately describable in our current language, which do not depend on human agency in any way.

ER₂ In order to have a language, we must presuppose that there is a world, independent of human representations, which is capable of being conceptualized in myriad ways.

⁵ One notable exception is Hund. See *op. cit.*

It is clear that these views are not necessarily incompatible, but it is equally clear that one might accept ER_2 while denying ER_1 . Indeed, the plausibility of ER_2 for the anti-realist will *depend* on it being distinct from ER_1 (i.e. on its being without content). The view that there are actual, articulable brute facts robs ER_2 of its plausibility for the anti-realist. It does so precisely because it begs the question *against* the anti-realist.

Consider: You claim that light is best explained as a wave. I retort that it is best explained as a particle. Someone else claims we should adopt the notion of a ‘wavicle’ in order to explicate light. We can all agree, regardless of our philosophical bearings, that there is an explanandum here—a set of data that we are trying to make sense of. Our disputes arise precisely because of the underdetermination of theory *by* data. No one need here deny that there is a world, independent of all of our representations, which we are trying to understand. This is what lurks at the core of ER_2 .

Disagreements arise between realist and anti-realist when we begin to assess divergent recommendations for understanding the world in a particular way. Whereas the realist wants to claim, in line with ER_1 , that there is *one* correct articulation of how things stand, the anti-realist doubts that there is some such definitive articulation. The anti-realist is sensitive to the fact that our descriptive vocabularies emerge in historical contexts, that our theories answer contingent questions, and that science and politics are never mutually exclusive. He thus doubts the truth of ER_1 . He doubts that the states-of-affairs that comprise the world are neatly articulable in a way that is independent of the circumstances (historical and otherwise) of the speaker.

So, as it stands, ER_2 , which Searle establishes via transcendental argumentation, is accurate. It is similarly pointless, however, as no one denies it. ER_1 , however, is a substantive view about the content of our ontological inventory. If Searle had shown that ER_1 was an accurate view of things, he would indeed have refuted the antirealist. But he has shown no such thing. ER_1 does not follow from ER_2 ; one cannot infer brute facts from the necessary postulation of a Kantian noumenal realm. If one could, Kant would never have needed to engage in his own Copernican Revolution, and the world would have managed to avoid Hegel.

To see precisely how egregious this philosophical maneuver is, consider the old adage that all facts are theory-laden. When Goodman claims that there are ‘ways of worldmaking,’ he need not deny that we presuppose some world external to our representation. In fact, he doesn’t deny this—he simply points out that there are ‘versions’ of this world which correspond to different systems of description. To offer any description of the world, then, is to ‘make’ the world into more than a dull substratum that we must presuppose in our everyday linguistic interaction. Of course, the systems of description we employ are not created *ex nihilo*. As Goodman puts it:

We start, on any occasion, with some old version or world that we have on hand and that we are stuck with until we have the determination and skill to remake it into a new one. Some of the felt stubbornness of fact is the grip of habit: our firm foundation is indeed stolid. Worldmaking begins with one version and ends with another.⁶

What Goodman is here intimating, and what we know well from Hanson, Kuhn, Cassirer, and Foucault (among others), is that systems of description, like predicates, can become entrenched in institutional life. It is precisely this entrenchment that gives particular utterances

⁶ P. 97, *Ways of Worldmaking* (New York: Hackett Publishing Company, 1978).

the appearance of rigidity, and that allows us to solve skeptical riddles like ‘Grue.’ Indeed, it is precisely the entrenchment of systems of description that makes sense of the term ‘paradigm’ in Kuhnian philosophy of science.

But this suggests that ER₁ is simply false, as the content of any so-called brute fact will depend on what system of description is the dominant one—on what predicates are entrenched in institutional life. Indeed, it is a merit of Berger’s sociology of knowledge that it can *explain* why some facts appear to be ‘brute.’ As Berger and Luckmann explain:

Reification is the apprehension of human phenomena as if they were things...The reified world is, by definition, a dehumanized world. It is experienced by man as a strange facticity, an *opus alienum* over which he has no control rather than as the *opus proprium* of his own productive activity (89).

This ‘reification’ occurs as much with recognizably social roles as it does with the institution of scientific practice. Because we have been socialized through a language, the basic elements of this language are seen as bits of a grand ontological inventory—as the brute pieces of an external world which we did not create. While this is a perfectly acceptable (and perhaps even necessary) product of socialization, it is nevertheless not the occasion to engage in high-church metaphysics. Collectively regarding some proposition as articulating a brute fact is as much a social phenomenon as is the buying and selling of goods with green paper.

While Searle is right that institutional facts are based on ER₂ (i.e. on the presupposition of an external world), he fails to see that the relationship can be regarded as exactly reversed for ER₁: substantive assertoric claims about the world are based on institutional facts about the meanings of terms. The assertoric speech act of the form ‘x is the case’ is based on a more primitive phenomenological fact: that we acknowledge constitutive rules of the form ‘x counts as y in c.’

Consider the following case: I make the non-trivial assertoric claim that ‘The morning star is the evening star.’ This is a claim about the empirical world—it asserts that a certain state-of-affairs is the case. It can be verified or falsified. In brief, it has all the flags of an assertoric speech act. Now, notice that this assertion depends on the two terms in the identifications having some significance. After all, if the claim *can* be verified or falsified, it *must* be the case that the claim *can be understood*. If I were to claim that ‘Degroner is in fact garap,’ no one could possibly verify what I had said, as no one could understand it. To make substantive claims about empirical reality requires that this reality is understood in approximately the same way by our interlocutors. But this does not show that any sort of robust realism is the case. In fact, it suggests rather the opposite: our assertoric claims about external reality depend crucially on our antecedent understanding of this reality. In this way, we might here speak of the primacy of phenomenology. Before I can make any substantive claims about the nature of some x, it must be the case that we collectively understand this x in a particular manner—that it is recognizable to us as an object of predication.

Given this line of argument, Searle can be accused of inverting the relation between constitutive rules and brute facts. Our everyday understanding of the world is littered with constitutive rules, and such rules are a condition for the possibility of making assertoric claims about empirical reality. Language itself suggests that ‘brute facts’ only have their basis in institutional reality. If language is constituted by constitutive rules which are the product of collective intentionality (as Searle claims), then what will count as the ‘correct’ description of extant phenomena will also depend on these rules. Using Nelson Goodman’s terms, we could put the point in an epigram: the more entrenched the predicate, the more brute the fact.

On this view, the analysis of institutional facts will not bottom out in brute facts, as brute facts can be shown to rely on institutional facts.

Concluding remarks

The above characterization of the relation between brute and institutional facts does not establish that the anti-realist position is a tenable one. It merely establishes that Searle's argument for ER has not shown that we must abandon anti-realism. Searle's social ontology does not establish that there are brute facts. Indeed, as I have hopefully made clear, that same ontology can be used to explain why we *regard* particular facts as brute, even when we should not. Given two divergent characterizations of the basis of brute facts, and little way to determine which is the more accurate, the lesson to be learned here might be that disputes about realism and antirealism ought to be abandoned—at least in our discussion of social ontology.

References

- Peter Berger and Thomas Luckman, *The Social Construction of Reality* (Garden City: Doubleday, 1966).
- Nelson Goodman, *Ways of Worldmaking* (New York: Hackett Publishing Company, 1978).
- John Hund, «Searle's *The Construction of Social Reality*,» *Philosophy of the Social Sciences*, 28(1), 1998, 122-131.
- John Searle, *The Construction of Social Reality* (New York: The Free Press, 1995).

J. Jeremy Wisnewski
East Carolina University
<jwisnewski@earthlink.net>

SORITES (ΣΩΠΙΤΗΣ), ISSN 1135-1349

<http://www.sorites.org>

Issue #16 — December 2005. Pp. 82-92

What is a Value Judgement?

Copyright © by SORITES and Georg Spielthener

WHAT IS A VALUE JUDGEMENT?

Georg Spielthener

The purpose of this paper is to clarify the concept of a value judgement. For ethical research a clear notion of value judgements is important for at least three reasons. Firstly, *moral* judgements are a certain kind of *value* judgements. Therefore, we can never get a clear concept of a moral judgement if we do not first know what a value judgement is. Secondly, if we do not know this, we cannot know how to *justify* moral judgements and therefore a well-founded *normative ethics* becomes impossible because this branch of ethical theory should enable us to justify our judgements about what things, acts, or qualities are good or bad, right or wrong. Thirdly, if there is no well-founded normative ethics, there is also no well-founded *applied ethics*, since the latter is at least partly the application of general normative theories to ethical problems.

In this paper I present my view on this problem, which is a version of non-descriptivism (or non-cognitivism) that is similar to, but not identical with, traditional non-descriptivist theories. The thesis I want to explain and argue for is that a person expresses his attitude towards something if and only if he makes a value judgement about this thing. In other words, I will defend the view that the expression of an attitude is a necessary and sufficient condition for making a value judgement. That is, if a person expresses in her judgement an attitude, she makes a value judgement and her judgement is not a value judgement if no attitude is expressed. Put more precisely, my thesis is that:

S makes a value judgement about *x* if and only if *S* expresses his attitude towards *x*.

The phrase ‘makes a value judgement’ is ambiguous. We can mean by it that we utter a value judgement, but obviously we can also make a judgement without expressing it. Usually, moral philosophers mean that we *utter* an evaluation when they say that we *make* a value judgement. In this paper I will keep to this tradition. It must, however, be emphasized that my account can also be applied to judgements that are not expressed. But this would require more space than I have here.

In what follows I will first explain this thesis by (I) clarifying the concept of an attitude, in (II) I defend the identity between having an attitude towards something and evaluating it, in (III) I distinguish value judgements from judgements that only seem to be evaluative, in (IV) I clarify what I mean by ‘expressing an attitude’, and in section (V) I give an argument for my view. I will, however, begin this paper by explaining in a very sketchy way what is novel about my account and why I think it is preferable to other meta-ethical theories. In short, my conception is an attempt to remedy the faults of rival theories while preserving their insights. To make this plausible, I will first deal with cognitivism.

Cognitivists (or, as I prefer to say, *descriptivists*) are partly right in many respects: Even though it is not correct that we make *always* an assertion about something if we make a value judgement about it, sometimes we do make an assertion if we evaluate something. Descriptivists are right when they emphasize that we can give logical reasons for our value judgements and not only explanatory reasons; they are also right when they claim that value judgements can be true or false; and when they are convinced that we can err when we make value judgements. My theory is in accordance with these insights, but in addition, it can also explain why all these claims of the descriptivists are only correct for some kinds of value judgements and not for all. More importantly, descriptivists are wrong in claiming that every evaluation of something is basically only a description of it and *nothing more*. I think the meta-ethical debate of the last century has sufficiently shown the inadequacy of this view. We have to distinguish between the evaluation of a thing and its description, but all versions of descriptivism are unable to make this distinction in a satisfactory way.

Since my account is a version a non-descriptivism, the difference to other theories of this kind is smaller but nevertheless considerable. (a) *Emotivism* was rightly criticised for its inability to distinguish between the causal problem of influencing attitudes and the logical problem of justifying our value judgements. In Stevenson's view, '[a]ny statement about any matter of fact which any speaker considers likely to alter attitudes may be adduced as a reason for or against an ethical judgement' (1944, p. 114). Obviously, this account fails to differentiate between explanatory and logical reasons. My conception does not blur this distinction. (b) According to R. Hare's *prescriptivism*, value judgements are necessarily prescriptive. A normative judgement is prescriptive if to subscribe to it is to be committed, on pain of being accused of insincerity, to doing this action, or, if someone else is required to do it, to willing that he does it. There are many problems connected with the prescriptive meaning of value judgements which I cannot discuss here. To my mind, Hare has not been successful in clarifying the prescriptive meaning of value judgements and more importantly, I think it is fair to say that his view that utilitarianism is the consequence of the universalizability and prescriptivity of ethical judgements is not convincing. I think I can give a clearer account of what value judgements are and how they can be justified.¹ (c) The difference to other kinds of *expressivism*² is less serious because my conception is itself a version of expressivism. I think, however, that I have explained more clearly than, for example, A. Gibbard or S. Blackburn what an attitude is and what it means to express an attitude. In addition, the other expressivists have (to my knowledge) neither distinguished between cognitive and non-cognitive attitudes nor between descriptive and non-descriptive value judgements and they have therefore not seen the relevance of these distinctions for the justification of value judgements. In short, I think I have developed a more detailed account of the nature of value judgements and their justification, but I have not dealt with the metaphysical and epistemological implications of my conception, as this has been done especially by Blackburn. Thus, my account and the investigations of these other expressivists are at least partly complementary. After these brief remarks about the relationship of my account to other meta-ethical theories, I will now explain my view of value judgements.

¹ I have dealt with the problem of justifying value judgements at length in Spielthener (2003, 238-59) and will do so again in a forthcoming paper.

² See e.g. Gibbard (1983; 1990; 1992) or Blackburn (1993; 1998).

I

The term ‘attitude’³ is also used in everyday discourse and refers there to a way of thinking (an example being when we say that a businessman does not have a professional attitude), but also to beliefs and convictions (when we say, for instance, that there is a change in public attitude towards abortion). However, in ethics ‘attitude’ is a technical term which differs in meaning from these everyday expressions. The essence of an ethical attitude is *being for or against* something. There are many things which are not indifferent to us. Most people are *against* wars, famines or torture and *for* freedom, peace and education. This means, they have a *con-attitude* (a negative attitude) towards the former things and a *pro-attitude* (a positive attitude) towards the latter ones. This general characterization can be supplemented by a negative explanation. A person has an attitude towards an object *if and only if* she is not indifferent to it. I am indifferent to the fact that the number 2 is a prime number. That I am indifferent to it means that I am neither *for* nor *against* it, and this means that I have no attitude towards this proposition.⁴ I am not indifferent to the fact that many people live in poverty and misery, and this means that I have an attitude towards it. The following two definitions can put more precisely what I have said so far:⁵

S has an attitude towards *x* if and only if *S* is for or against *x*.

S has an attitude towards *x* if and only if *S* is not indifferent to *x*.

These definitions of ‘attitude’ are very general, broader than the definitions usually given in the social sciences,⁶ but in philosophy as well.⁷ It is, however, this concept most ethicists refer to when they use it⁸ or analyse it,⁹ and it includes exactly those elements that are essential for the ‘evaluative attitudes’ that are relevant to ethics.

So far I have defined attitudes as *being for or against* something. It seems, however, that I must explain this definition as well. What does it mean to be *for* or *against* something? In

³ Since it is important to distinguish the *use* of a term from its *mention*, I will use single quotation marks when I mention a term.

⁴ Stevenson (1944, 4, 13) says in this context that we give only a ‘detached’ or ‘disinterested description’ of something. As Broad (1985, 48) points out, rightly, we must distinguish between two kinds of indifference, namely (a) the ‘balanced indifference’ which we have when our pro- and con-attitudes to a thing cancel each other out, and (b) an ‘uninterested indifference’ which we have when we have neither a positive nor a negative attitude to an object. As already said, I mean by ‘indifferent’ this second kind of indifference.

⁵ These definitions are precisising (or reforming) definitions. This means, they are intended to make the meaning of terms, which are already used in ordinary language, more precise to make them herewith more useful. But such definitions do not give a completely new meaning to the definiendum.

⁶ Compare, for example, Rockeach (1973).

⁷ For Edwards (1955, 29), the concept of an attitude comprises the disposition to encourage others to act accordingly, which is a much narrower concept.

⁸ For example Stevenson (1944; 1963), Nowell-Smith (1957), Perry (1926; 1954). Perry uses the term ‘interest’, but he refers by it to the same concept. In a similar way Dreier (1997, 87) refers to this concept by using ‘desire’. Many philosophers use the term ‘desire’ (e.g. Broad, 1985), but it is more restricted than ‘attitude’ and therefore less suitable. For example, we cannot desire something that happened in the past, but we can have an attitude to it – and past events are ethically important.

⁹ For example Pitcher (1958).

philosophy and psychology it is usual to explain this as a *disposition*.¹⁰ Dispositions are said to be not directly observable tendencies to react in a given situation in a specific way. Traditionally, psychologists have investigated three groups of such reactions that manifest attitudes. (a) *Cognitive* reactions are, for example, our beliefs in something; (b) *affective* reactions are emotions towards the object of an attitude; and (c) *conative* reactions which are the tendencies of persons to act in a certain way in relation to the object of the attitude. For example, when a person has a pro-attitude towards quiet flats (i.e. is *for* such flats), she will show this *cognitively*, for instance, in her wondering how to get such a flat, *conatively* in her being motivated to rent one, and *affectively* by her happiness when she has found such a flat. The idea is that *being for or against* something is not, or, at least, not necessarily, a conscious state, but a tendency to react in a certain way. Only these manifestations are said to be conscious and observable at least introspectively.

However, I will not attempt here to define the meaning of being for or against something by referring to the concept of a disposition because this concept is burdened with many problems and therefore the definiens would be more unclear than the definiendum. These problems are not only of a philosophical nature (logical, epistemological, and semantic), but there are also empirical problems. Even if a series of psychological investigations show that attitudes are related to cognitive, affective, and conative reactions,¹¹ they have failed to specify them in a way which could significantly improve our understanding of attitudes. In particular, the relation between being for or against something and actions is still relatively undetermined because it is obvious that actions are not only influenced by attitudes but also by a number of other factors.¹² Therefore, I will not analyse the concept of *being for or against* something any further, but will treat it as a basic concept. This seems to me no problem because I think every adult person knows what it means to be for or against something. There is thus no need for a further analysis of this concept. Similar to pleasure and pain, it seems that being for or against something cannot be analysed satisfactorily, but, as with pleasure and pain, we are so acquainted with them that there is also no need for such an analysis.¹³

However, some distinctions will be useful to prevent misunderstandings. It is obvious that there are many different ways of being for or against something. When I *wish* that something is the case, I am, in a way, for it, but I am also for it, when I *want* it. But wishing and wanting are not the same. Similarly, I am against something when I *despise* it, but also when I *disapprove* of it. Nevertheless we have to distinguish between these concepts. Already these few examples can suggest that there are *different kinds* of attitudes. I think this is correct and

¹⁰ Compare e.g. Eagly and Chaiken (1993, 1-2), Fishbein and Ajzen (1975, 8), Stevenson (1944, 60) or Urmson (1968, 41) who writes that somebody can have an attitude to a religion 'whether he thinks about golf or mathematics, whether he is asleep or awake, whether he feels gay, angry, tired or bored.' A dispositional explanation for the narrower concept 'desire' is also preferred by Smith (1987). But there are also philosophers who seem to see dispositions as *conscious* states, e.g. Richardson (1997, 52) who characterises 'desire' as a 'psychological state' and not as a psychological *trait*.

¹¹ See Eagly and Chaiken (1993, 6-7).

¹² Compare, e.g. Fishbein and Ajzen (1975). For an overview about the problem of how attitudes are related to actions see e.g. Eagly and Chaiken (1993, 155-218).

¹³ Broad (1967) writes about pleasure and pain saying that they are something 'which we cannot define but are perfectly well acquainted with ...' (p. 22). It seems to me that the same is true for being for or against something.

a number of philosophers have been of the same opinion.¹⁴ However, if there are these kinds of attitudes, then ‘attitude’ is a generic term. This would imply, firstly, that there is some feature which is common to all attitudes and, secondly, that there are distinguishing features (a *differentia specifica*) which, when added to or combined with the generic feature, mark off the different kinds of attitudes. I think both are true. The feature which is common to all attitudes is *being for or against something*. Whenever we approve or disapprove of something, find something beautiful or ugly, good or bad, we are either for or against it. To determine the distinguishing features is a difficult task. Urmson (1968, pp. 48-61) and Pitcher (1958) have tried to find them for *approvals*, Swinburne (1985) for *desires*, and Anscombe (1963) has analysed the concept of *wanting*. I am not convinced that they have given adequate analyses of these attitudes, but their investigations, and also those of many others, show that it is possible to determine the distinguishing features of different attitudes. Philosophers have analysed many kinds of (evaluative) attitudes. All of them have in common that they are ways of being for or against something.¹⁵ Even though it is very desirable to determine also the distinguishing features of different attitudes, for our present purpose we need only the common feature of all attitudes, namely that they are ways of being for or against something.

II

Having explained what I mean by ‘attitude’, I will now defend the equivalence between having an attitude towards something and evaluating it. In short, my thesis is that:

S evaluates *x* if and only if *S* has an attitude towards *x*.

I have argued extensively for this thesis of the equivalence between ‘evaluating a thing *x*’ and ‘having an attitude towards *x*’.¹⁶ In this paper I give only one argument which may also be useful to illustrate what I mean by evaluating something and having an attitude towards it. The argument starts from the notion of evaluating something. I will try to show that to evaluate something means to be for or against it, or in other words, not to be indifferent to it.

Let us take four different situations: (a) If we consider states of affairs and events which we clearly and undoubtedly evaluate, for instance, pain, wars, torture, but also happiness, health, or joy, then we will see that we are not indifferent to them. We are either *for* or *against* them. When I find a piece of music beautiful, an action mean, or a medical treatment painful, then I am not indifferent to these things. It is, for example, paradoxical to say, ‘Beethoven’s ninth symphony is beautiful, but I don’t like it,’ if we mean ‘beautiful’ really in an evaluative sense. (b) On the other hand, we can consider cases that are not evaluative. Years ago I had to teach students whose knowledge of history was very poor. But since I knew that this was the result of their difficult living conditions, I did not blame them. I did not evaluate the students negatively and so I did not have a con-attitude towards them. (c) But there are situations where we are clearly for or against something. When we learn about child abuse, torture, or abductions (especially when we are personally involved in such cases), we

¹⁴ Compare e.g. Brandt (1996, 57-8; 1998), Davidson (1985, 20, 129, 151; 1985a, 151-2), Nowell-Smith (1957, 175), Stevenson (1944, 90; 1963, 2, 223-4) and Urmson (1968, 43-4). Of course, these kinds of attitudes can then be classified, e.g. as ‘egoistic’ or ‘altruistic’ as proposed by Stevenson (1963, 5).

¹⁵ Stevenson (1944) mentions e.g. ‘purposes’, ‘aspirations’, ‘wants’, ‘preferences’, ‘desires’, ‘interests’, ‘approving’, ‘favour’, ‘ideals’ and ‘aims’ – together with their contraries. Clarke (1985, 43) lists ‘desire’, ‘wish’, ‘hope’ and ‘be interested in’, and Nowell-Smith (1957, 99) gives also a list of attitudes. Besides the examples listed here, we could add e.g. ‘recommendation’, ‘attraction’, ‘admiration’, ‘esteem’ and ‘indignation’, together with their respective antonyms. But every list is necessarily incomplete because in every ordinary language new terms emerge which can express attitudes.

¹⁶ See Spielthener (2003, 3-73).

are not indifferent. We are against such incidents. At the same time we evaluate them. It is not the case that we are against such events but do not evaluate them. But this means that we evaluate something if we are for or against it. (d) Finally, let us consider cases which are indifferent for us. If I am indifferent to the order of my books on the shelf, to the current weather in Patagonia, or to the question whether the number 284 is divisible by 3, then I am not evaluating these things. In this case I do *not* judge, for instance, that Marx's books *must not* be placed next to the Canon Law, that the weather in Patagonia *should* be better or that it is *bad* that 284 is not divisible by the number three.

Already these few simple examples seem to me sufficient to justify my claim that the nature of valuing lies in being for or against something – or to formulate it with greater precision, that a subject *S* evaluates an object *x* if and only if *S* is for or against *x* (or is not indifferent to *x*). In any case, this result follows from the points (a) and (c) and from (b) and (d) as well. That is, if it is correct what I have said in (a) and (c) or in (b) and (d), then I have shown that evaluating something is tantamount to being for or against it because this result follows logically from (a) and (c) and also from (b) and (d).

In the light of what has been said so far, the following argument for the correctness of my thesis of the identity of attitudes and evaluation can be given. As pointed out, a person has an attitude towards an object if and only if she is for or against it, and we are for or against something exactly when we evaluate it. From these two premises follows that a person evaluates a thing exactly when she has an attitude towards it. With greater precision this argument can be put as follows: (1) *S* has an attitude towards *x* iff *S* is for or against *x*. (2) *S* evaluates *x* iff *S* is for or against *x*. (3) Therefore, *S* evaluates *x* if and only if *S* has an attitude towards *x*. Premise (1) is the in section (I) presented and defended definition of attitudes. Premise (2) is the result of this section's explanation of the concept of an evaluation. The conclusion (3) is the thesis of the identity between attitudes and evaluations, which follows logically from the two premises. This means, if the premises are correct, the conclusion must be so as well and I am justified in using these terms interchangeably.

III

Since the main thesis of this paper says that we express an attitude towards something if we make a value judgement about it, judgements are often not value judgements even if they seem to be because they contain words like 'good', 'should', or 'must'. We are often tempted to regard such judgements as value judgements; nevertheless they are sometimes at best *pseudo* value judgements. Let us consider some kinds of them. (a) *Reporting 'value judgements'* report an evaluation but they are not evaluative themselves. If someone says, 'You should go to church on Sundays,' then it can be that he is only *reporting* an obligation. Such judgements seem to be value judgements. On closer examination, however, it turns out that they are statements about value judgements. They state that it is a duty to go to church on Sundays, which is a (true or false) statement, but not a value judgement. Another kind of reporting 'value judgements' are some judgements about *conventions*. If someone says, for instance, that men must stand up when they greet a lady, he may only be reporting a convention without supporting it. In this case, however, he is not making a value judgement. What is reported can also be *standards*. Urmson (1968, pp. 62-71) criticises the view that value judgements express attitudes by emphasizing that persons may only use standards when they assert an evaluative judgement. In this case they would not express an attitude. For example, a shop assistant in a grocery might use a standard for cheese when he says, 'This is excellent cheese,' without expressing herewith his attitude towards the cheese. Urmson is correct in stressing this point, but he has overlooked the fact that in this case this judgement is not a value judgement, but a statement claiming that the cheese meets the standards for excellent cheese. (b) Also *hypocritical 'value judgements'* are in my view not value judgements. When

a husband says that he should be faithful to his wife, without attaching any importance to faithfulness, then he is not making a value judgement – even if, on first sight, he seems to make (c) Also many ‘value judgements’ in the *third person* are de facto not value judgements. ‘John won’t think that’s a good idea,’ need not be a value judgement. I can only express the (true or false) belief that John will not like it. However, judgements in the third person can be evaluative. If a child says, ‘Mother has forbidden me to nibble between meals,’ she can utter a value judgement. She does so, if she expresses an attitude in uttering this judgement.¹⁷ Even though we tend to regard such *pseudo* value judgements as genuine value judgements, they are not evaluative. Not every judgement which looks like a value judgement on the surface is one on closer examination, and often judgements which seem to be pure statements are in fact value judgements. Whether or not a judgement is a value judgement cannot easily be seen.

IV

My thesis claims that value judgements express attitudes. It is this view which makes my theory ‘expressivistic’,¹⁸ and therefore I must explain what I mean by ‘expressing an attitude’.

That we express attitudes when we make value judgements is a view taken by many. The question, however, is what this metaphor means. To start with, it is the judging person who expresses her attitude. Some authors are at least misleading when they write that *judgements* or *concepts* express attitudes.¹⁹ Evaluations are expressed by persons (and eventually by other entities) and they express them by making judgements. (Please note that ‘making judgements’ must be understood as explained in the introduction.) Other philosophers think that *terms* are expressive.²⁰ But together with Brandt (1996) I take the view that only judgements can be expressive, strictly speaking. I can express my attitude by saying, ‘that was cruel,’ but only in specific contexts by ‘cruel’ alone, and in those cases this word is a judgement itself.

Even though many authors take the view that value judgements express attitudes, only a few have seen the need to analyse this concept. One of these few is Smart (1984), who holds that the assertion of a sentence expresses an attitude if we (the recipients) can probabilistically conclude that the judging person has an attitude (see p. 40), and he attaches great importance

¹⁷ Cf. Stevenson (1963, 127). From these examples we must distinguish *ironic* value judgements. They *are* value judgements, but through the irony the evaluation is reversed. If someone says ironically, ‘This was really a good dinner,’ he *is* evaluating the dinner, but in a negative way and not, as the word ‘good’ suggests, a positive way. (To this also compare Hare, 1952, 160 and Stevenson, 1944, 83).

¹⁸ As the formulation of the thesis shows, I consider here only the question of when something is a value judgement for the *judging* person. Other authors (especially Stevenson, 1944; 1963) have also investigated the question when something is a value judgement for its *recipients*. Of course, there is a difference. Something can be a value judgement for the judging person, but not for the recipient, and vice versa. For a salesclerk in a music shop the judgement ‘This is an interpretation of Schubert’s Impromptus by Radu Lupu’ may be a value judgement, for a buyer, who is not very well acquainted with piano music and its interpreters, it can be a simple statement; on the other hand, the information that the CD with the Impromptus costs Euros 20 may be a statement for the seller but can be (economically) evaluative for the buyer because he thinks it is too expensive. However, my restriction to the question when something is a value judgement for the person who is judging is justified because it is this question which is essential for ethics.

¹⁹ Compare, e.g. Stevenson (1963, 207). Of course, we must first *have* an attitude, only then can we *express* it. It seems that some authors (e.g. Audi, 1991, 96) think that we can express attitudes without having them. In my view this is impossible. We can *pretend* to express an attitude (or an actor can play the expression of an attitude), but we can only *express* e.g. our contempt when we *have* it.

²⁰ E.g. Edwards (1955) or Ayer (1954).

to his view that not sentences but only their *assertions* express attitudes.²¹ But this explanation is unsatisfactory because a person can express her attitude towards something even though we have no reason to conclude that she has an attitude towards it. This happens often when we are dealing with foreign cultures or when we are confronted with languages which we speak only badly. But it can even happen when we make a judgement such as, ‘Mr. *X* didn’t congratulate Ms. *Y* on her birthday.’ This can be a value judgement (*X* can be blamed for not doing so), but we need not have a reason to infer an attitude of the judging person. Therefore, it is necessary to have a closer look at this concept.

To understand the meaning of ‘expressing an attitude’, we must first distinguish between the *description* and the *expression* of an attitude. If someone says that he disapproves of the death penalty, then he is making a statement *about* his attitude. What he says is true if he really disapproves of it and false if he does not do so. If a person claims that the death penalty should be abolished (and is really committed to her judgement), then she is *expressing* an attitude, but does not claim that she has one. It was this simple but fundamental distinction between expressing and describing attitudes, which distinguishes *emotivism* (and other versions of non-descriptivism) from the so-called *ethical subjectivism* (a version of descriptivism or cognitivism).²² For ethical subjectivism value judgements are statements *about* attitudes (*descriptions of them*), for emotivistic and expressivistic theories they are *expressions of attitudes* (which, however, does not necessarily mean that they are *only* this).

To clarify what it means to express an attitude, I will begin with an example that can illustrate the essential elements of my analysis. Let us take the judgement, ‘The Jupiter Symphony was composed by Mozart.’ This *can* be a value judgement and I will here assume that it is one. Let us examine the implications of this judgement. (1) I believe that Mozart composed the symphony. That is, I have a belief about the object of my evaluation. (2) I assert in my judgement what I believe about this object (namely that this symphony was composed by Mozart). (3) I have an attitude to the Jupiter Symphony. If I am indifferent to it, my judgement is not a value judgement. When I say, for instance, ‘The piano sonata K. 457 was written by Mozart,’ but do not have an attitude towards this sonata, I do not utter a value judgement, but only a statement. (4) I have also an attitude towards what I believe about the Jupiter Symphony, namely that Mozart composed it. My judgement, ‘The Jupiter Symphony is in C major,’ is not a value judgement, even if I have an attitude towards this symphony, but not to its key. (5) I have my attitude towards this symphony *because* I have an attitude towards its composer. If I had an attitude towards the Jupiter Symphony only because it is in C major and not because it was composed by Mozart, my judgement, ‘The Jupiter Symphony was composed by Mozart’, would not be a value judgement, even if I have an attitude towards this symphony.

This example suggests the following analysis of the concept ‘expressing an attitude’:

S expresses in his or her judgement about *x* an attitude towards *x*, if (1) *S* has his or her attitude towards *x* at least partly (a) because *S* has an attitude towards something that *S* believes about *x* or (b) because this attitude was caused by non-cognitive factors; and (2) what is believed about *x* or this non-cognitively caused attitude is (a) asserted by the judgement or (b) the reason for the chosen expression (often, but not always, on the basis of linguistic conventions).

²¹ We may deduce that someone is fond of symbolic logic from the fact of his or her assertion that symbolic logic is good stuff, but not from ‘Symbolic logic is good stuff’ itself (Smart, 1984, 10).

²² See, e.g. Ayer (1952), Edwards (1955), Stevenson (1944), or any introduction into meta-ethics.

Some examples of value judgements will be useful to illustrate this analysis. Mr. *A*, a pious Christian, is against abortion and expresses his disapproval by saying, ‘Abortion is an offence against God’s will.’ The object of his evaluation (*what* he evaluates) is abortion (or the proposition that abortions are carried out). *A* has an attitude towards abortions, and he has it at least partly because he believes that they are against God’s will, to which he has an attitude as well (he disapproves of acting against God’s will). In addition, he is stating what he believes. (This example satisfies the conditions 1a and 2a.)

However, Mr. *A* could express his disapproval in a different way, for example, by judging, ‘Abortion is a sin.’ Also in this case he has a negative attitude because he believes that abortion is against God’s will (and he has a con-attitude towards this as well), but he does not assert that. What he believes is the reason why he calls abortion a sin. (The example satisfies the conditions 1a and 2b.)

Let us take a third case. Ms. *B* tells her friend, ‘I despise greed for money.’ Let us assume that this is a value judgement (undoubtedly, it can be one but it is not necessarily one). In this case *B* has an attitude towards greed for money (she despises it). *Ex hypothesi*, she has this attitude because of her upbringing (we assume that she has acquired this attitude through operant conditioning).²³ This means, Ms. *B* is not against greed for money because she believes that avarice has certain qualities but because of some purely psychological processes; and this means, in other words, that her attitude was caused by non-cognitive factors. She asserts in her judgement that she has an attitude (in saying that she despises greed for money) – which, however, does not preclude that she also expresses this attitude. (This example satisfies the conditions 1b and 2a.)

Finally, let us consider a fourth case. *A* groans, his face contorted with pain, ‘Damn headache!’ This is a value judgement if *A* expresses an attitude by uttering it. We assume also here that this attitude was not caused by cognitive factors, that is, we assume that it is *non-cognitive*. However, *A* does not state that he has this attitude. The attitude is the reason for the chosen linguistic expression ‘damn headache’. *A* could have chosen other ways of expressing his pain. The way we express our attitudes is to a large extent determined by linguistic convention. For example, the word ‘cool’ has been a term used by many to express their pro-attitudes. But not all expressions of attitudes are in this way determined by conventions. ‘He is a foreigner’ can express an attitude, but in my opinion there is no linguistic convention according to which ‘foreigner’ is a term which expresses attitudes.²⁴ (This fourth example satisfies the conditions 1b and 2b.)

V

So far I have explained what I mean by the thesis that expressing an attitude is sufficient and necessary for making a value judgement. In this section I will put forward an argument for it. Elsewhere I have given a number of reasons,²⁵ but here I must confine myself to one succinct argument.

One can argue for my thesis in a logically valid way as follows: (1) *S* makes a value judgement about *x* iff *S* expresses an evaluation of *x*. (2) *S* evaluates *x* iff *S* has an attitude towards *x*. (3) If *S* evaluates *x* exactly when *S* has an attitude towards *x*, then *S* expresses her

²³ I have discussed such acquisition of attitudes in Spielthener (2003, 109-24).

²⁴ Compare Urmson (1968, 24-37).

²⁵ See Spielthener (2003, 42-47).

evaluation of x exactly when S expresses her attitude towards x . It follows from this (4) that S makes a value judgement about x if and only if S expresses her attitude towards x .

Explanation: The first premise makes an uncontroversial assertion about value judgements. It is a truism that we evaluate something if we make a value judgement about it. The second premise is the thesis of the equivalence between evaluating and having an attitude. I have given one argument for it in section II (but more elsewhere). The third premise seems to me unproblematic and self-evident. If a person evaluates a thing exactly when she has an attitude towards it (as premise 2 says), then she will *express* her attitude towards this thing exactly when she expresses her evaluation of it. If having an attitude towards x is tantamount to evaluating it, then expressing an attitude towards x is tantamount to expressing an evaluation of it. The conclusion (4) follows logically from the premises (1) – (3). This conclusion, however, is the main thesis of this paper, namely that we make a value judgement about something if and only if we express an attitude towards it.

References

- Anscombe, G. E. M. (1963). *Intention* (2nd ed.). Oxford: Blackwell.
- Audi, R. (1991). *Practical reasoning*. London-New York: Routledge.
- Ayer, A. J. (1952). *Language, truth, and logic*. New York: Dover.
- Ayer, A. J. (1954). *Philosophical essays*. London: Macmillan.
- Blackburn, S. (1993). *Essays in quasi-realism*. New York-Oxford: Oxford University Press.
- Blackburn, S. (1998). *Ruling passions: A theory of practical reasoning*. Oxford: Clarendon Press.
- Brandt, R. B. (1996). *Facts, values, and morality*. Cambridge: Cambridge University Press.
- Brandt, R. B. (1998). The rational criticism of preferences. In C. Fehige & U. Wessels (Eds.), *Preferences* (pp. 63-77). Berlin-New York: de Gruyter.
- Broad, C. D. (1967). *Five types of ethical theory* (6th ed.). London: Routledge & Kegan Paul.
- Broad, C. D. (1985). *Ethics*. Dordrecht: Martinus Nijhoff.
- Clarke, D. S. (1985). *Practical Inferences*. London: Routledge & Kegan Paul.
- Cooper, N. (1981). *The diversity of moral thinking*. Oxford: Clarendon.
- Copp, C. (1995). *Morality, normatively, and society*. New York: Oxford University Press.
- Davidson, D. (1985). Handlungen, Gründe und Ursachen. In D. Davidson, *Handlung und Ereignis* (S. 19-42). Frankfurt am Main. Suhrkamp. (First published 1963: Journal of Philosophy)
- Davidson, D. (1985a). Beabsichtigen. In D. Davidson, *Handlung und Ereignis* (S. 125-152). Frankfurt am Main. Suhrkamp. (First published 1978)
- Dreier, J. (1997). Humean doubts about the practical justification of morality. In G. Cullity & B. Gaut (Eds.), *Ethics and practical reason* (81-99). Oxford: Clarendon Press.
- Eagly, A. H. & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth: Hartcourt Brace.
- Edwards, P. (1955). *The logic of moral discourse*. New York: The Free Press.
- Fishbein, M. & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading/Mass.: Addison-Wesley.

- Gibbard, A. (1983). A nonkognitivist analysis of rationality in action. *Social Theory and Practice*, 9.
- Gibbard, A. (1990). *Wise choices, apt feelings. A theory of normative judgment*. Cambridge/Mass.: Harvard University Press.
- Gibbard, A. (1992). Moral concepts: Substance and sentiment. In J. E. Tomberlin (Ed.), *Philosophical perspectives (Vol. 6, pp. 199-221)*. Atascadero: Ridgeview Publishing Company.
- Hare, R. M. (1952). *The language of morals*. Oxford: Oxford University Press.
- Milo, R. D. (1984). *Immorality*. Princeton: Princeton University Press.
- Nowell-Smith, P. H. (1957). *Ethics*. Oxford: Blackwell.
- Ossowska, M. (1980). *Moral norms: A tentative systematisation*. Amsterdam: North-Holland.
- Perry, R. B. (1926). *General theory of value*. New York: Longmans Green and Co.
- Perry, R. B. (1954). *Realms of value. A critique of human civilization*. Cambridge/Mass.: Harvard University Press.
- Pitcher, G. (1958). On approval. *Philosophical Review*, 67, 195-211.
- Richardson, H. S. (1997). *Practical reasoning about final ends*. Cambridge: Cambridge University Press.
- Smart, J. J. C. (1984). *Ethics, persuasion and truth*. London: Routledge & Kegan Paul.
- Smith, M. (1987). The Humean theory of motivation, *Mind*, 96, 36-61.
- Spielthener, G. (2003). *Grundlagen der Ethik*. Frankfurt/Main: Peter Lang.
- Stevenson, C. L. (1944). *Ethics and language*. New Haven and London: Yale University Press.
- Stevenson, C. L. (1963). *Facts and values. Studies in ethical analysis*. New Haven: Yale Univ. Press.
- Stofer, H. (1955). *Über das ethische Werturteil*. Basel: Verlag fuer Recht und Gesellschaft.
- Swinburne, R. (1985). Desire. *Philosophy*, 60, 429-445.
- Urmson, J. O. (1968). *The emotive theory of ethics*. London: Hutchinson.

Georg Spielthener
University of Zambia
<georg_spielthener@yahoo.de>

HYPER LIBERTARIANISM AND MORAL LUCK

Gerald K. Harrison

I

Although the debate is not yet settled, many now accept that Frankfurt-style cases, or something very like them, refute the principle of alternate possibilities:

PAP: A person is morally responsible for what he has done only if he could have done otherwise.

In what follows I will assume that PAP has been refuted. The object of this paper is to investigate the prospects for libertarianism in a post-PAP world. The results of my investigation are surprising. Contrary to received opinion, it will turn out to be rather a good thing for libertarianism if PAP goes.

I will argue that the kind of libertarian positions which survive post PAP are ones which have the resources to make sense of moral luck in a way unavailable to the compatibilist. This, I claim, provides a non-question begging reason to favour libertarianism over compatibilism. If I am right then the refutation of PAP, far from heralding a new renaissance for compatibilism, might in fact be the beginning of the end for that position.

II

Historically PAP has provided the main support for libertarian positions on freedom and moral responsibility. Although compatibilists could argue for a conditional analysis of PAP according to which it can be met even if determinism is true, this was never a particularly natural reading of the principle. So, PAP gave the libertarians an advantage over compatibilists.

However, if PAP is false then things are different, or so it is assumed. Although there is logical room for libertarian positions post PAP, the advantage seems to lie with the compatibilist, if it lies with anyone. This is a picture I want to question, but for the time being we need to get clear over why this picture has come to be received wisdom.

First we need to get clear that the refutation of PAP does not directly refute libertarianism. With PAP out of the picture the question becomes whether causal determinism in the actual sequence rules out moral responsibility. There is room here to insist that it does, and hence there is room for a variety of what Fischer has referred to as hyper libertarian positions (Fischer 1999, pp. 129-130fn).¹ The term is meant to be disparaging, but I rather like it and will use it hereafter. A hyper libertarian then, is simply a libertarian who thinks that determinism undermines moral responsibility for reasons that are not to do with the availability of alternative possibilities — not, in other words, to do with an agent's ability to do otherwise.

¹ Amongst contemporary hyper-libertarians can be numbered Randolph Clarke; Eleanor Stump and David Hunt.

There are a number of reasons why the hyper libertarian might claim that determinism rules out moral responsibility. They might claim that causal determinism in the actual sequence is incompatible with our being the initiators or originators of our decisions and choices (where this is not a requirement of having a certain kind of control, but a separate requirement in its own right), or that it might rule out an essential type of creativity. As a hyper libertarian myself, what I claim is that determinism deprives one of ownership over one's decisions. We need to be the ultimate or, to use a phrase of Kane's «buck-stopping» explanation of our decisions in order for them to be truly ours (see Kane 1989 p. 254). This we cannot be if determinism is true, for the explanation of why we made one decision rather than another will always trace back to factors external to our deliberative process. Kane refers to this requirement as a requirement for ultimacy (Kane 1989, p. 254). Again, it is important to recognise that this concern is not about control. As far as I am concerned the only type of control worth wanting (or indeed, the only kind of control that is intelligible) is of a sort compatible with determinism. But important though control is, it is not all that there is to freedom: in addition to controlling our decisions we need also to own them in the sense that I have just outlined. According to my position then, it is ownership that is threatened by determinism.²

But a compatibilist is simply going to object that it begs the question against the compatibilist to insist that indeterminism is needed for ownership, creativity, importance, significance and so forth. (see Fischer 2003 pp. 198-210). After all, there are perfectly good compatibilist interpretations of ownership, creativity, importance, significance and any other plausible notion that the libertarian might care to maintain is threatened by determinism. The compatibilist could say something like this: your decisions and choices are significant in as full a sense as anyone could want because the world had to go through you to get that to happen. So, where the libertarian stresses independence from the world, the compatibilist could stress the agent's indispensability from the world and so on (on this see Fischer 2003, p. 208). More could be said here, (Fischer, for example, has his own, compatibilist, account of ownership) but the basic point is clear. Some reason needs to be given why we should favour libertarian interpretations of ownership, independence, creativity and so forth over compatibilist ones. As Fischer says «once the debate is shifted away from the relationship between causal determinism and alternative possibilities, it is difficult to present a non-question begging reason why causal determinism rules out moral responsibility» (Fischer 2003, p. 201). Indeed, compatibilism should be favoured «insofar as our basic views about ourselves — our views of ourselves as persons and as morally responsible — should not be held hostage to the discoveries of a consortium of scientists about the precise nature of the equations that describe the universe» (Fischer 2003, p. 211).

So, it looks as if the libertarian is indeed disadvantaged by the passing of PAP. My own hyper libertarian position survives the falsity of PAP, but according to Fischer and others it does not have anything to recommend it to any who are not already converted to libertarianism, and may in fact look like the less attractive option when we bear in mind that compatibilism does not hold certain of our views «hostage to the discoveries of a consortium of scientists». If you like, we have nothing to lose and everything to gain by being compatibilists in a post PAP world.

² Of course, there are, amongst hyper libertarians, those who would want to supplement this kind of account with an account of agent-causation. I think that this is unnecessary and unhelpful and returns the issue to being one of control. This was the problem with the old debate. PAP led to all the focus being on control. Now, I think that if we can possibly help it, it would be better if we can sell libertarianism without having to make recourse to agent-causation, simply because the resulting theory will not require any previous commitment to libertarianism.

III

In fact, there is a very good reason to favour the kind of hyper libertarianism I have outlined. To see why, we need first to understand just why the old-style libertarian was so concerned about an ability to do otherwise. To answer this question we need to get clear about just what compatibilist control amounts to.

Consider that, according to a popular compatibilist account of control what it is for a decision or choice to have been controlled is simply for it to have been the output of a certain type of mechanism: one sensitive or responsive to an appropriate range of inputs. The kind of control mechanism that is relevant to moral responsibility is one that is reason-responsive (the term ‘reason-responsiveness’ was coined by Fischer in his *Metaphysics of Free Will* and he and Ravizza have further developed and refined this conception of control in their book *Responsibility and Control: A Theory of Moral Responsibility* 1998). In other words, the mechanism that issues in the agent’s decisions and choices must be one that is sensitive or responsive to a certain degree (for otherwise it will fail to be a control mechanism) and what it is sensitive or responsive to must be reasons. Other things will plausibly be important too, for instance the history of the mechanism in question, and the way in which the mechanism came to issue in the decision or choice that it did — but strictly speaking these kinds of concern are not to do with control. So, having control of a kind relevant to moral responsibility means simply having a mechanism of a certain sort, sufficient to give you the «power to be moved by reasons» (Dennett 1984, pp. 18-19; p. 25; pp. 50-51; p. 98). Essentially then, what having control amounts to is having capacities and dispositions to respond in certain ways to a relevant range of inputs.

This account of compatibilist control is far too brief, but it would be impossible to say more within the scope of this paper, suffice it to say that this paper is addressed to those compatibilists who accept something like the above account of control.

Libertarians have concerns about this type of control. It is not that they deny that it is an account of a type of control. After all, no libertarian should deny that we can distinguish between controlled and uncontrolled events even if determinism turns out to be true. Rather, the concern is over the adequacy of this type of control when it comes to moral responsibility. This concern has been nicely captured recently by Paul Russell:

On the face of it, therefore, the agent is liable to blame and retribution, on the moderate reasons-responsive account, merely for possessing capacities that he is not able to exercise control over. [à] in the case of [wrongdoing] the agent appears to be simply unlucky enough to be moderately reasons-responsive and placed in circumstances where the mechanism fails to track the reasons that were present (2002, p. 595)

I believe that it is plausible that it is this type of concern that lay behind the commitment to PAP. The thought was that if an agent has only compatibilist control over their decision and choices, then they nevertheless lack control over how their mechanism operates in the actual sequence. Given that they are, by hypothesis, morally responsible for the decisions that they make, this means that the agent will be exposed to a certain kind of moral luck. It is their bad luck, for instance, that they possess a mechanism which, in these exact circumstances, will issue in this, morally reprehensible decision. For in different, but relevantly similar circumstances, it would have issued in a different, blameless, or even praiseworthy decision.

By relevantly similar circumstances I mean any possible world in which both the mechanism and the morally relevant reasons are held fixed. Unless the mechanism in question is, to use Fischer’s term, strongly reason responsive, then there will always be such possible worlds — possible worlds in which the mechanism and the morally relevant reasons are held fixed, and where the mechanism delivers a different output for which the agent is, by hypothesis, morally responsible (Fischer and Ravizza 1998, pp. 41-42)

In response to this concern one move would be to insist that strong reason-responsiveness is what is required for moral responsibility. If an agent is strongly reason responsive then they are not exposed to the kind of moral luck in question. Whatever decision the strongly responsive mechanism delivers, it would have delivered the same decision in all scenarios in which both it and the morally relevant reasons are held fixed. It is therefore no matter of luck that it delivered the decision that it did in the actual sequence. However, strong reason responsiveness is far too demanding a control requirement which would in practice all but rule out moral responsibility. Indeed, to avoid exposure to this kind of moral luck, the reasons to which the mechanism would have to be strongly sensitive would have to be the moral reasons present, which would have the upshot that no agent could ever do wrong culpably.

To see how it is this kind of concern about moral luck that lay behind PAP consider that if the agent is not to be exposed to moral luck of the kind in question (and assuming that strong reason responsiveness is an implausible demand for moral responsibility) she will need some additional type of control — a type of control that is not captured by talk of reason-responsive mechanisms of various kinds. Unlike mechanism accounts, this type of control will require genuinely alternative possibilities. In the actual circumstances in which the mechanism operated and failed to track reason, it needs to have been genuinely possible for it to have succeeded in tracking reason. Only if such a genuine possibility exists, can there be room for the agent to have had additional control over how the mechanism operated in the actual sequence. For only if there are genuinely available alternative possibilities does the agent have the ability to track, or fail to track reason in the actual sequence. So, it is a concern about moral luck which gives rise to a requirement for a kind of executive control which can only be accommodated if there are genuinely available alternative possibilities.

To reiterate then, if we only have compatibilist control, then it seems fundamentally unfair to hold us morally responsible for the decisions and choices that we make in the actual sequence, for we would have made different decisions in alternative, but relevantly identical sequences. The kind of executive control associated with PAP would have ruled out this kind of moral luck.

IV

With PAP gone, then the kind of executive control in question is not available or not actually required for moral responsibility. But just because PAP is false, this does not mean that the concerns about moral luck were not legitimate, and nor does it mean that they magically disappear. This concern about moral luck does not seem to me to be question-begging in any interesting way. I would say that the kind of moral luck in question is *prima facie* problematic, and in the absence of an adequate explanation it has to count as a serious weakness in the compatibilist's position that they must affirm its existence. But of course this same concern now applies to my own hyper libertarian position, for I acknowledge that the kind of control outlined above is all the control required for moral responsibility.

So the situation is this, post-PAP both the hyper-libertarian and the compatibilist should agree about control. The only kind of control relevant to moral responsibility is the mechanism kind outlined above. However, the original concerns about the adequacy of compatibilist control are still with us. Both hyper-libertarian and compatibilist must affirm the reality of the kind of moral luck outlined. Yet, the reality of this kind of moral luck is *prima facie* problematic.

It is now that my hyper libertarian position can be seen to have an advantage over compatibilism. For my position can make sense of moral luck in a way unavailable to the compatibilist. Consider that the kind of moral luck in question can, we have seen, be ruled out if the agent is strongly reason responsive. If, in other words, they had full compatibilist

control. It would, as I have already said, be implausible to insist upon strong reason responsiveness as a condition of moral responsibility. But nevertheless what we can notice is that for a compatibilist the strongly reason-responsive agent could, in principle, be morally responsible for the decisions and choices that they make. The same is not true for anyone holding the type of hyper libertarian position that I have outlined. For the only way in which an agent could be strongly-reason responsive is if their mechanism is internally deterministic. Only if the mechanism is internally deterministic will we be able to say of it that it would have delivered the same decision in all relevant possible worlds — worlds in which the same morally relevant reasons are present and the mechanism is held fixed. If the mechanism were internally indeterministic then there would always be some possible worlds in which the same reasons are present and the mechanism is held fixed, yet a different decision issues. Internal indeterminism therefore entails something less than strong reason responsiveness. If internal indeterminism is required, as I maintain that it is, then something less than strong reason responsiveness is actually a requirement of moral responsibility. This means that the kind of freedom needed for moral responsibility actually exposes the agent to the moral luck in question. Exposure to moral luck is a condition of having the kind of freedom necessary for moral responsibility.

Here we can see the truth in something that the compatibilists have often claimed. Namely, that indeterminism does nothing to enhance control and will in fact diminish it (Kane 2003, p. 318). The libertarians used to charge that this was question begging. But now that, post PAP, hyper libertarian and compatibilist agree about control, they should agree to this claim. In the absence of any special kind of libertarian control, making an agent's reason responsive mechanism internally indeterministic will diminish the degree of control it can be said to deliver, for it will invariably make it less responsive than it would otherwise have been.

It is important to recognise here that if the agent's reason-responsive mechanism is internally indeterministic, that does not preclude its being sufficiently reason-responsive for moral responsibility. My target audience here are those compatibilists who allow that indeterminism does not rule out moral responsibility. I take it that the appeal of the reason-responsive account of control is precisely that it has this upshot. So long as only some degree of compatibilist control is required — so, just some degree of reason responsiveness - then an internally indeterministic mechanism can, in principle, be sufficiently responsive to satisfy the control requirement to any of my target compatibilist's satisfaction. Note also, that any compatibilist who held otherwise would then lose the compatibilist their supposed advantage over hyper libertarianism — it would be to place our «views of ourselves as persons and as morally responsible [in] hostage to the discoveries of a consortium of scientists» (Fischer 2003, p. 211).

So, moral luck begins to make sense when we take the hyper libertarian's perspective. No such account is available to the compatibilist. They cannot insist upon internal indeterminism as a condition of moral responsibility without giving up their compatibilism. So they are stuck with having to simply insist that moral luck is a brute fact unamenable to explanation. By contrast, the hyper libertarian such as myself can offer a principled explanation of the phenomena in question. The compatibilist is challenged to present a systematic account of moral luck, or else concede that there is a powerful, non-question begging reason to prefer libertarianism to their own position.

So, the concern about moral luck stems from focussing only on the control condition associated with moral responsibility. But once we understand that control is not all that there is to freedom there opens up the possibility that some other condition of freedom might be in tension with (though not contradict) the control condition. Ownership, according to the

libertarian position I have outlined, is in tension with the control condition insofar as it can only be satisfied if the agent has less than the kind of full control that would rule out moral luck. Hence moral luck is explicable as being the cost of freedom. Freedom, when properly understood, involves being exposed to the kind of moral luck in question, for it requires that we have less than the kind of full control that would rule out such exposure.

The unfairness of moral luck remains. But we have seen that to rule out unfairness completely would require a level of control incompatible with being free. And fundamentally an agent is morally responsible not because it would be fair to hold them morally responsible, but because they made their decision freely. It is an agent's freedom which explains and makes comprehensible their moral responsibility: in short, moral responsibility is about freedom, and not about fairness.

V

I shall now try to address what I take to be some of the more pressing objections to some of the claims I have made above. I should add however, that what I have said is supposed only to show how one kind of moral luck could be intelligible given our received ideas about moral responsibility. It is not supposed to be a comprehensive solution to the larger problem presented by other kinds of moral luck. I believe it can go some way towards such an account because I believe that the kind of moral luck we have been talking about is the most fundamental, but that is a matter for another occasion. On that front what I have said is merely suggestive.

To some criticisms, then. Central to my argument is the claim that indeterminism internal to the operation of the agent's reason responsive mechanism would rule out strong reason responsiveness. In other words, introducing indeterminism in the way needed to secure ownership would diminish control, and would rule out complete or full control. But this claim might be challenged. Mele has suggested a way in which the libertarian's demand for internal indeterminism could be met without sacrificing any compatibilist control.

[I]f it is causally undetermined whether a certain belief will enter into Jones's deliberation, then Jones lacks [compatibilist control] over whether the belief enters his deliberation. But this need not be an impediment to Jones's control over how he deliberates in light of the beliefs that do enter his deliberation. (Mele 1995, p. 215)

But here I need to make clear in what sense I meant that indeterminism needs to be internal to the deliberative process. For I would claim that an agent's beliefs form part of their circumstances, and are as such external, not internal to the deliberative process. They are, if you like, at the input end, whereas the indeterminism needs to be at the output end. It needs to be undetermined what decision the process will issue in at the point of decision making. Following Kane, the indeterminism needs to be at the point of decision making. The reason is that if the explanation of the agent's decision traces back to factors over which the agent did not exercise control then the agent is not morally responsible for that decision. For if this is the case then the agent is controlled by the past in a way that precludes ultimacy and ownership. Mele's proposal violates the ownership condition, for by hypothesis it is not under Jones's control what beliefs enter his deliberation. Things are different if it is indeterministic what decision will issue at the point of decision making. For by hypothesis the explanation cannot now trace back to factors not under the agent's control, because the decision is the output of the control mechanism.

So, I would claim that Mele's proposal does not introduce indeterminism in a way that would meet the requirements of ownership. It is only if the indeterminism is internal to the control mechanism in such a way that it must diminish control, that the ownership condition is met. However, I am willing to concede that there may be variations of hyper libertarian

position according to which indeterminism need only enter in the way outlined by Mele. The precise strategic location for indeterminism can be a matter of debate amongst libertarians. But whilst such positions may be coherent, they would be incapable of offering an account of moral luck and thereby would offer no advantage over compatibilism. Indeed, the compatibilist position would, for reasons given earlier, actually have a slight advantage over such positions. So, the claim that the indeterminism needed for ownership will diminish compatibilist-control is, I hold, defensible, and recommends itself over libertarian alternatives, precisely because it facilitates an account of moral luck that would otherwise be unavailable.

A second kind of objection can now be raised. What I have said above about the decision being literally undetermined until the moment of choice might lend itself to being interpreted as the claim, associated with some libertarians, that the decision in question must be uncaused by prior events. This in turn will then lead to the objection that the resulting decision will be insufficiently connected to the agent's reasons for action.

However, from the fact that it needs to have been indeterministic, up to the moment of decision, does not mean that the decision in question will not have been antecedently caused. My claim is that ownership or ultimacy is achieved only if the ultimate explanation of why one decision was made rather than another stops with the agent. But this is not the same as the claim that the explanation of why the decision made was made needs to stop with the agent. The difference here is between a plain and a contrastive explanation. A plain explanation is one that answers the question «why P». A contrastive explanation answers the question «why P rather than Q». If universal indeterminism obtains so that every event is indeterministically caused then we will be unable to give contrastive explanations. If it is undetermined whether T will cause P to occur or Q to occur, then we cannot explain why P rather than Q occurred in the actual sequence. Citing T will be of no help. However, it does not follow that we will be unable to give plain explanations. We will still be able to explain why P occurred, for we can cite T. T plainly explains P. But T does not contrastively explain why P rather than Q. Consider that if determinism obtained, then it is in principle possible that T would both plainly explain P, and contrastively explain why P rather than Q. So determinism obtaining will mean that what previously would only plainly explain, can now contrastively explain. It is the absence of the relevant contrastive explanation that is required by ownership, not the absence of a relevant plain explanation.

So, if the agent's decision making process is indeterministic up to the moment of decision, that does not mean that we cannot give a plain explanation of the agent's decision in terms of their antecedent causes. So there is no reason to think that we will not be able to link the agent's decision with their reasons. What it does mean however, is that we cannot give a contrastive explanation of why the decision making process resulted in that decision rather than a different one.

Now this gives rise to another, third objection. If, in the actual sequence, the agent makes decision P, then we can cite the agent's prior reasons in explanation of why this particular decision was made. But if we roll back time to just before the decision was made, and run things through again, then the decision making process might issue in a different decision — decision Q. Given that the prior reasons will be the same, how can Q be explained in terms of the agent's prior reasons?

This means that, if Jane is deliberating about whether to vacation in Hawaii or Colorado, and gradually comes to favour and choose Hawaii, she might have chosen otherwise (chosen Colorado), given exactly the same deliberation up to the moment of choice that in fact led her to favour and choose Hawaii (exactly the same thoughts, reasonings, beliefs, desires, dispositions, and other characteristics — not a sliver of difference). (Kane 2003, p. 302).

Here I follow Kane in claiming that the relevant deliberative processes are ones where we have opposing reasons — we have reasons in favour of making one decision and reasons in favour of making an alternative decision.³ So, along with Kane, I claim that the kind of indeterminism needed for ownership can arise only on those occasions where the agent's deliberations take the form of trying to sort out which amongst competing reasons to act from or on. These occasions provide us with examples where the deliberative process can be indeterministic insofar as it is undetermined how this deliberative process will turn out, but at the same time, however it turns out, the decision in question will have been made on the basis of reasons. I will here let Kane, to whose account I am clearly heavily indebted, offer some clarification.

Imagine that the businesswoman is trying or making an effort to solve two cognitive problems at once, or to complete two competing (deliberative) tasks at once — to make a moral choice and to make a choice for her ambitions. With respect to each task she is being thwarted in her attempt to do what she is trying to do by indeterminism. But in her case, the indeterminism does not have a mere external source; it is coming from her own will, from her desire to do the opposite. Recall that the two crossing neural networks involved are connected, so that the indeterminism which is making it uncertain that she will do the moral thing is coming from her desire to do the opposite, and vice versa. She may therefore fail to do what she is trying to do. But I argue that, if she nevertheless succeeds, then she can be held responsible because, like them, she will have succeeded in doing what she was trying to do. And the interesting thing is that this will be true of her, whichever choice is made, because she was trying to make both choices and one is going to succeed. (2003, pp. 312-313)

This account of internal indeterminism is quite consistent with the mechanism in question being sufficiently reason responsive to satisfy any plausible compatibilist-control condition. Obviously, a mechanism that is responsive or sensitive to competing reasons such that it is indeterminate how it will respond, cannot deliver full compatibilist control. But that was always the point. So a reason-responsive mechanism that is responsive to competing reasons, can nevertheless deliver a high enough degree of responsiveness to qualify for moral responsibility where control is concerned, and it will also deliver the kind of internal indeterminism needed to ensure ownership of the decisions that issue. For we will be unable to explain why the agent acted for the reasons that they did, rather than the competing reasons. The agent here terminates the explanation and thereby achieves the kind of significance and ultimacy required for ownership according to my account.

Finally, a fourth objection I can anticipate regards PAP. I have been assuming that PAP is false, yet the account of hyper libertarian freedom that I have given is one that says that, in fact, at the point of decision making it was undetermined what decision the agent would make. This was required in order to rule out the relevant contrastive explanation and thereby satisfy the ultimacy or ownership condition. But then it would seem that ultimacy or ownership does require that there are alternative decisions that, in the actual sequence, it was possible for the agent to have made. It was just such a possibility that Frankfurt-style cases were designed to rule out. If we are accepting that PAP is false on the basis of Frankfurt-style cases, then we cannot at the same time make it a condition of moral responsibility that there exist these kinds of alternative possibilities.

³ This will obviously not always be the case. But then sometimes, maybe even quite often, it will be the case that it is not indeterministic prior to decision what decision the agent will make. In such cases the resulting decision will either be one for which the agent cannot bear moral responsibility, or will be one for which they can bear responsibility because what has made the decision inevitable is the result of earlier, genuinely free choices on the part of the agent.

Two responses. Firstly, many compatibilists accept that Frankfurt-style cases cannot be constructed in which all alternative possibilities are ruled out, for to do this would require assuming determinism, which just begs the question against the libertarian (see Fischer 1999 p 122). Rather, what Frankfurt-style cases do, and why they refute PAP, is they rule out relevant alternative possibilities, which is to say alternative possibilities that the agent can «access». I have argued elsewhere that even if there is an alternative possibility in which the agent responsibly decides otherwise, Frankfurt-style cases can still show that the agent did not have the ability to decide otherwise, because the agent did not have the right kind of access to the alternative in question (its obtaining depending upon something lucky or improbable happening). To make this argument in any detail would require a separate paper, so I will say no more about this point apart from to reiterate that I believe (in line with many others) that Frankfurt-style cases refute PAP, but not by ruling out the possibility of deciding otherwise, only by ruling out the ability to decide otherwise. Since it was securing the ability to do otherwise that alternative possibilities were originally thought necessary, nothing I have argued here implies that the refutation of PAP is unstable.

But, just in case this does not satisfy those who might make this criticism, let us accept, for the sake of argument, that Frankfurt-style cases do rule out alternative possibilities in which the agent responsibly decides otherwise (after all, some have tried to design Frankfurt-style cases in which all relevant alternative possibilities are rule out — see Hunt, D. 2000; Stump 1999). The point about ultimacy and ownership is that these are conditions that concern the origin, or way that the decision came about in the actual sequence. The agent is still the ultimate contrastive explanation where they are the indeterministic cause of their making the decision that they did, even if in the actual sequence there were external constraints meaning that no other decision would have been possible, in those precise circumstances. For the point is that whilst prior conditions might fully explain why the decision in question occurred (because of the presence of some suitably placed counterfactual intervention device) they cannot fully explain why the agent made the decision in question for the reasons that they did. So, even on the hypothesis that such scenarios are constructable without begging any questions (and I believe that they are not) , it is still the case that indeterminism, strategically located, can secure the kind of ultimacy needed for ownership. The point would be that ultimacy does not actually demand that no contrastive explanation can be given of why the decision was made, but rather that no relevant contrastive explanation can be given — one making reference to the agent's reasons.

Conclusion

The contemporary debate over the compatibility of determinism and moral responsibility has paid insufficient attention to the problem of moral luck. This is a major oversight if, as I believe, much of the concern over the threat from determinism is traceable to concerns over moral luck. I have argued here that once we focus on this issue, we will see that in a post-PAP world both libertarian and compatibilist will find themselves committed to affirming the reality of moral luck. However, hyper libertarianism has the advantage here, as it has the resources to make sense of the reality of moral luck in a way unavailable to the compatibilist.

In the introduction to the latest edition of his collection on Free Will Watson says, «[s]uppose that we are causal systems whose operations are highly probable, but not certain, given their antecedents ..[t]his supposition satisfies the incompatibilist requirement, but it hardly gives us what we are after» (2003, p. 9). On the contrary, what I have argued is that it does give us what we are after. Once PAP is out of the way, we can start to see how the space opened up by an indeterministic physical process is a space needed for ownership, not for control. This requirement is one that should recommend itself to the compatibilist too, because it provides something that both sides need, namely a systematic explanation of the

reality and comprehensibility of moral luck. So, by focussing our attention on the real problems associated with compatibilist control, the refutation of PAP helps us to see that we have a very powerful, non-question begging reason to favour a hyper libertarian position over compatibilist alternatives.

Bibliography

- Clarke, R. 1995: «Indeterminism and Control» *American Philosophical Quarterly* 32 pp. 125-138.
- Fischer, J. 2003: «Frankfurt-style Compatibilism» in *Free Will* ed. Watson, G. New York: Oxford University Press.
- Fischer, J. 1999: «Recent Work on Moral Responsibility» *Ethics* 110 pp. 93-139.
- Fischer, J. & Ravizza, M. 1998: *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
- Haji, I. 1998: *Moral Appraisability*. New York: Oxford University Press.
- Hunt, D. 2000: «Moral Responsibility and Unavoidable Action», *Philosophical Studies* 97.2, pp. 195-227.
- Kane, R. 2003: «Responsibility, Luck, and Chance: Reflections on Free Will and Indeterminism» in *Free Will* ed. Watson, G. New York: Oxford University Press.
- Kane, R. 1996: *The Significance of Free Will*. New York: Oxford University Press.
- Kane, R. 1989: «Two Kinds of Incompatibilism», *Philosophy and Phenomenological Research* 50 pp. 219-254.
- Mele, A. 1995: *Autonomous Agents: From Self-Control to Autonomy*. New York: Oxford University Press.
- Stump, E. 1996: «Libertarian Freedom and the Principle of Alternative Possibilities,» in *Faith, Freedom, and Rationality*, ed. Jordan, J. and Howard-Snyder, D. Lanham, MD: Rowman & Littlefield.
- Pereboom, D. 1995: «Determinism al Dente,» *Noûs* 29 pp. 21-45. York: Oxford University Press.
- Watson, G. 2003: «Introduction» in *Free Will* ed. Watson, G. New York: Oxford University Press.
- Russell, P. 2002: Review of *Responsibility and Control: A Theory of Moral Responsibility* in *Canadian Journal of Philosophy*, vol 32 no. 4 pp. 587-606.

Gerald K. Harrison
University of Durham
U.K.

<g.k.harrison@durham.ac.uk>

SORITES
An Electronic Quarterly of Analytical Philosophy
ISSN 1135-1349
COPYRIGHT NOTICE AND LEGAL DISCLAIMER
© 1996, 2004 The SORITES Team

- (01) **SORITES** is not in the public domain. It is the intellectual property of the **SORITES** team. In accordance with Spanish and international Law, all issues of **SORITES** are Copyright-protected throughout the Planet.
- (02) Each issue of **SORITES** taken as a whole belongs to the electronic publisher, namely the SORITES team.
- (03) However, the individual papers published in **SORITES** remain the intellectual property of their respective authors, with only these two restrictions: (i) No part of any such paper may be printed or displayed elsewhere, or incorporated into a publication of any sort, unless **SORITES** is therein clearly and explicitly mentioned as the primary source; and (ii) The authors agree to be bound by the other terms and provisions contained in this Copyright Notice, esp. to grant readers the entitlements warranted by clauses (06), (08), (09), and (19), and to abide by the electronic publisher's claims and commitments set forth in clauses (14) and (15).
- (04) The authors of the included papers and the electronic publisher, the **SORITES** team — whether jointly or separately, as the case may be — hereby reserve all rights not expressly granted to other parts in this Copyright Notice.
- (05) In compliance with Spanish Law, this issue of **SORITES** has been legally registered, three diskette-copies being deposited with the competent authorities, namely the «Depósito Legal» office of the Autonomous Community of Madrid. (Legal Deposit Registration: M 14867-1995.)
- (06) A license is hereby granted without fee for anybody to freely make as many unmodified copies as they wish of this issue of **SORITES IN ITS INTEGRITY**, give such copies to anyone, and distribute this issue of **SORITES** via electronic means or as printed copies, **PROVIDED** no part thereof is altered or omitted, and especially **NEITHER THIS COPYRIGHT NOTICE NOR ANY OF THE COPYRIGHT BOXES ON TOP OF THE PAPERS IS REMOVED, AMENDED, OR OBSCURED**.
- (07) In this context, *the issue of **SORITES** as a whole* is meant to consist in: either (i) a single file (be it its official version as a WordPerfect 5.1 document or any unofficial version released by the **SORITES** team as an undivided file); or (ii) a collection of files produced by slicing one of the entire-file versions in order to facilitate handling, browsing or downloading. In the latter case, the conveyor is bound to distribute the whole collection. (In this context printed copies of this issue of **SORITES** are taken to be equivalent to electronic copies, their distribution being subject to the same conditions.)
- (08) This issue of **SORITES** may be sold for profit or incorporated into any commercial material only with the previous explicit consent granted by the SORITES team. Otherwise, no fee may be charged for its circulation. An exception is granted to non-profit organizations, which are hereby authorized to charge a small fee for materials, handling, postage, and general overhead.
- (09) Private copying of single papers by any lawful means is allowed when done in good faith and for a fair use, namely for purposes of teaching, study, criticism or review.
- (10)
- [i] No copy or partial copy of a detached paper may be publicly distributed or circulated without its author's acquiescence.
 - [ii] The SORITES team hereby allows any such distribution, subject to the author's consent, provided the copyright box on top of the paper is literally incorporated into the distributed copy or partial copy.
 - [iii] The phrase 'copy or partial copy' covers any print-out, extended quotation, electronic posting or display as well as any other reproduction.

- (11)
- [i] No part of this issue of **SORITES** may be delivered to a plurality of individuals — whether in writing or by any other means — unless the source is clearly and explicitly acknowledged.
 - [ii] In particular, no part of this issue of **SORITES** or of any paper therein included may be conveyed to others by means of reproduction, quotation, copy or paraphrase, without a clear and explicit acknowledgement of the issue of **SORITES**, the author's name and the paper's full title.
 - [iii] Whenever the quotation occurs within a publication, it is also mandatory to mention the date of publication and the official pages (as shown within the Copyright box on top of the paper), as well as the ISSN (1135-1349) and the official website: <http://www.sorites.org>.
- (12) Any use of copies or partial copies of this issue of **SORITES**, or of papers therein included, not explicitly permitted in this Notice, will be a violation of the authors' and the electronic publisher's intellectual property; perpetration of, or complicity with, any such violation will be regarded as forgery or plagiarism — which, besides being, in any case a civil tort, may be a crime under current legislation.
- (13) This issue of **SORITES** is provided «as is», without any guarantee of any kind. The electronic publisher, the **SORITES** team, disclaims all warranties, whether expressed or implied, including, without limitation, the implied warranties of fitness for any particular purpose with respect to the papers included in this issue. By furnishing this document, the **SORITES** team does not grant any license or endorses any commitment except in so much as explicitly set forth in the present Copyright Notice.
- (14) The electronic publisher, the **SORITES** team, does not necessarily agree with the authors' views or arguments. The electronic publisher cannot certify the accuracy of any quotations or references contained in the papers.
- (15) Each author vouches alone for the originality of the papers they submit to **SORITES** and for their compliance with established Copyright laws. Acceptance of a manuscript is done in good faith under the assumption the originality claim is truthful. The electronic publisher — i.e. the **SORITES** team — does not pledge itself for the accuracy of such declarations.
- (16) Neither the **SORITES** team nor the authors are liable for any real or imaginary damages suffered as a result of downloading, reading, using, quoting or circulating any materials included in this issue of **SORITES**. The user assumes, at their own risk, full responsibility for the proper use of this issue of **SORITES**.
- (17) Downloading, reading or in any other way using this issue of **SORITES** or any part thereof entails full acceptance of the hereinabove stated terms and provisions. If, after downloading a file containing this issue of **SORITES** or a part thereof, a user fails to agree to the conditions contained in this Notice, they must discontinue using the material and irrecoverably erase or destroy the downloaded file, so as not to occasion any third-part's unfair use thereof.
- (18) Although, thanks to a permission kindly granted by the system's administrators, *SORITES* is displayed at the internet host <www.ifs.csic.es/sorites/> belonging to the Spanish research institution CSIC, the magazine is not sponsored or endorsed by the CSIC, the only owner and publisher being the **SORITES** team, under the auspices of no academic establishment.
- (19) A specific license is hereby granted for this issue of **SORITES** to be freely mirrored at any website, provided all conditions stated above are fully adhered to. No previous approval of the **SORITES** team is required for such a display.

Madrid. April 10, 1995

(Updated: December 31, 1996; April 27 1998; October 3, 2002; December 31, 2004)

The **SORITES** Team

RELEASE NOTICE

This issue of *SORITES* is made available in several formats, but its only official version is the WordPerfect 5.1 document released with filename:

sorite16.wp

Although each version, whether official or not — as initially released by the **SORITES** team — is an entire seamless file, it may be splitted down into chunks in order to facilitate downloading, browsing, transferring or e-mailing. In such cases, the unity of this issue of **SORITES** as a whole must be preserved by keeping the ensuing collection intact.