The Metaphysics of Consciousness

Royal Institute of Philosophy Supplement: 67

Edited by Pierfrancesco Basile, Julian Kiverstein & Pauline Phemister

Royal Institute of Philosophy Supplements

The Royal Institute of Philosophy Supplements are published twice a year. Institutional subscribers to the journal *Philosophy* receive the supplements as part of their subscription. The following supplements are available to purchase as books.

Volume 37	1994	*Philosophy, Psychology and Psychiatry*	(ISBN 0521469023)
38	1995	Philosophy and Technology	(ISBN 0521558166)
39	1995	Karl Popper: Philosophy and Problems	(ISBN 0521558158)
40	1996	Philosophy and Pluralism	(ISBN 0521567505)
41	1996	*Verstehen and Human Understanding*	(ISBN 0521587425)
42	1997	Thought and Language	(ISBN 0521587417)
43	1998	Current Issues in Philosophy of Mind	(ISBN 0521639271)
44	1999	German Philosophy since Kant	(ISBN 0521667828)
45	1999	Philosophy and Public Affairs	(ISBN 0521667844)
46	2000	Logic, Cause & Action	(ISBN 0521785103)
47	2000	Philosophy, the Good, the True and the Beautiful	(ISBN 0521785111)
48	2001	Philosophy at the New Millennium	(ISBN 0521005086)
49	2001	Naturalism, Evolution and Mind	(ISBN 0521003733)
50	2002	Time, Reality & Experience	(ISBN 0521529670)
51	2002	Logic, Thought and Language	(ISBN 0521529662)
52	2003	Philosophy and the Emotions	(ISBN 0521537347)
53	2003	Minds and Persons	(ISBN 0521537339)
54	2004	Modern Moral Philosophy	(ISBN 0521603269)
55	2004	Agency and Action	(ISBN 0521603560)
56	2005	Philosophy, Biology and Life	(ISBN 0521678455)
57	2005	The Philosophy of Need	(ISBN 0521678447)
58	2006	Political Philosophy	(ISBN 0521695597)
59	2006	Preferences and Well-Being	(ISBN 0521695589)
60	2007	Narrative and Understanding Persons	(ISBN 9780521714099)
61	2007	Philosophy of Science	(ISBN 9780521718967)
62	2008	Being: Developments in Contemporary Metaphysics	(ISBN 9780521735445)
63	2008	Kant and Philosophy of Science Today	(ISBN 9780521748513)
64	2009	Epistemology	(ISBN 9780521138581)
65	2009	Conceptions of Philosophy	(ISBN 9780521138574)
66	2010	Philosophy as Therapeia	(ISBN 9780521165150)

From Volume 13 onwards the Series is published by Cambridge University Press and some earlier titles are also available.

This journal issue has been printed on FSC-certified paper and cover board. FSC is an independent, non-governmental, not-for-profit organization established to promote the responsible management of the world's forests. Please see www.fsc.org for information.

The Metaphysics of Consciousness

ROYAL INSTITUTE OF PHILOSOPHY SUPPLEMENT: 67

EDITED BY

Pierfrancesco Basile
Julian Kiverstein
Pauline Phemister

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, CB2 1RP,
United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 8RU, United Kingdom
32 Avenue of the Americas, New York, NY 10013–2473, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

Printed in the United Kingdom at the University Press, Cambridge
Typeset by Techset Composition Ltd, Salisbury, UK

A catalogue record for this book is available from the British Library

ISBN 9780521173919
ISSN 1358-2461

Contents

Notes on Editors and Contributors v

Preface 1

Sprigge's Ontology of Consciousness 5
LEEMON McHENRY

How to Become Unconscious 21
STEPHEN R. L. CLARK

The Road to Substance Dualism 45
GEOFFREY MADELL

Fundamental Singleness: How to Turn the 2nd Paralogism into a Valid Argument 61
GALEN STRAWSON

It Must be True – But How Can it Be? Some Remarks on Panpsychism and Mental Composition 93
PIERFRANCESCO BASILE

Phenomenal Holism 113
BARRY DAINTON

Is There a Metaphysics of Consciousness Without a Phenomenology of Consciousness? Some Thoughts Derived from Husserl's Philosophical Phenomenology 141
EDUARD MARBACH

Making Sense of Phenomenal Unity: An Intentionalist Account of Temporal Experience 155
JULIAN KIVERSTEIN

Time in Consciousness, Consciousness in Time 183
DAVID COCKBURN

Quality, Thought and Consciousness 203
HOWARD ROBINSON

Concessionary Dualism and Physicalism 217
WILLIAM SEAGER

The Representational vs. the Relational View of Visual Experience 239
BRIAN P. McLAUGHLIN

Consciousness: Don't Give Up on the Brain 263
KENNETH AIZAWA

Index of Names 285

Notes on Editors and Contributors

KENNETH AIZAWA

Kenneth Aizawa is the Charles T. Beaird Professor of Philosophy at Centenary College of Louisiana. He is the author of *The Systematicity Arguments* (2003) and, with Fred Adams, *The Bounds of Cognition* (2008).

PIERFRANCESCO BASILE

Pierfrancesco Basile teaches Philosophy at the University of Bern, Switzerland. His publications include two books, *Experience and Relations: An Examination of F. H. Bradley's Conception of Reality* (1999) and *Leibniz, Whitehead and the Metaphysics of Causation* (2009), as well as several essays on issues in metaphysics, process thought, British idealism and early analytic philosophy.

STEPHEN R. L. CLARK

Stephen R. L. Clark was Professor of Philosophy at the University of Liverpool. His books include *The Nature of the Beast* (1982), *Animals and their Moral Standing* (1997), *Biology and Christian Ethics* (2006), *G. K. Chesterton: Thinking Backwards, Looking Forwards* (2006) and *Understanding Faith: Religious Belief and Its Place in Society* (2009).

DAVID COCKBURN

David Cockburn is Professor of Philosophy at the University of Wales, Lampeter. His publications include *Other Human Beings* (1990), *Other Times: Philosophical Perspectives on Past, Present and Future* (1997), and *An Introduction to the Philosophy of Mind* (2001). He edited the Royal Institute of Philosophy volume on *Human Beings* (1991), and has published a wide range of papers on themes in philosophy of mind, ethics, Wittgenstein, philosophy of religion, and philosophy of time.

BARRY DAINTON

Barry Dainton is Professor of Philosophy at the University of Liverpool. His publications include *Stream of Consciousness* (2000, 2nd edition 2006), *Time and Space* (2001, 2nd edition 2010), and *The Phenomenal Self* (2008).

JULIAN KIVERSTEIN

Julian Kiverstein is a Teaching Fellow in Philosophy at the University of Edinburgh. The author of a number of papers on consciousness, temporality and the self, he is also editing *Heidegger and Cognitive Science* (2010, with Michael Wheeler) and *Decomposing the Will* (2010, with Tillmann Vierkant and Andy Clark).

GEOFFREY MADELL

Geoffrey Madell was Senior Lecturer in Philosophy at Edinburgh University. He is the author of *The Identity of the Self* (1981), *Mind and Materialism* (1988) and *Philosophy, Music and Emotion* (2002). He has published numerous articles on the philosophy of mind, particularly on the issues of personal identity and the mind-body problem.

EDUARD MARBACH

Eduard Marbach was Professor of Philosophy at the University of Bern, Switzerland. His publications include *Das Problem des Ich in der Phänomenologie Husserls* (1974), *Mental Representation and Consciousness: Towards a Phenomenological Theory of Representation and Reference* (1993), and *An Introduction to Husserlian Phenomenology* (1993, co-authored with Rudolf Bernet and Iso Kern). He has edited Husserl's *Phantasie, Bildbewusstsein, Erinnerung. Zur Phänomenologie der anschaulichen Vergegenwärtigungen* (1980, *Husserliana*, Vol. 23) and is the author of several papers on phenomenology and the philosophy of mind.

LEEMON McHENRY

Leemon McHenry teaches Philosophy at California State University, Northridge. His books include *Whitehead and Bradley: A Comparative Analysis* (1992), *Consciousness, Reality and Value: Essays in Honour of T. L. S. Sprigge* (2007, co-edited with Pierfrancesco Basile), and *British Philosophers, 1800–2000* (2002, co-edited with Philip Dematteis and Peter Fosl). He is the author of numerous articles and book chapters on topics in metaphysics, philosophy of science and medical ethics.

BRIAN P. McLAUGHLIN

Brian P. McLaughlin is Professor II in the Philosophy Department at Rutgers University and a member of the Executive Council of the Rutgers University Cognitive

Science Center. He has published numerous articles in a wide range of fields in philosophy. Recently, he was a co-editor of *Contemporary Debates in the Philosophy of Mind* (2007) and a co-editor of *The Oxford Handbook of Philosophy of Mind* (2009).

PAULINE PHEMISTER

Pauline Phemister is Reader in Philosophy at the University of Edinburgh. Her publications include *The Rationalists: Descartes, Spinoza and Leibniz* (2006) and *Leibniz and the Natural World: Activity, Passivity and Corporeal Substances in the Philosophy of Leibniz* (2005). She has published widely on topics in early modern philosophy, especially in the areas of metaphysics, ethics, and mind-body relations.

HOWARD ROBINSON

Howard Robinson is University Professor at Central European University, Budapest. He is author of *Matter and Sense* (1982, re-issued in paperback 2009) and *Perception* (1994), editor of the World's Classics edition of Berkeley's '*Principles*' *and* '*Three Dialogues*', and of various recent articles and book chapters on the philosophy of perception, the mind-body problem and idealism.

WILLIAM SEAGER

William Seager is Professor of Philosophy at the University of Toronto. He has published many articles in the philosophy of mind and two books on consciousness, *Metaphysics of Consciousness* (1991) and *Theories of Consciousness* (1999).

GALEN STRAWSON

Galen Strawson teaches in the Philosophy Department at the University of Reading and is a regular visitor at the Philosophy Program at the City University of New York Graduate Center. He is the author of *Selves: An Essay in Revisionary Metaphysics* (2009), *Real Materialism and Other Essays* (2008), *Mental Reality* (1994, new edition 2009), *The Secret Connexion: Realism, Causation and David Hume* (1989), and *Freedom and Belief* (1986, new edition 2010).

Preface

In a recent article entitled 'The Mind-Body Problem at Century's Turn',[1] one of the leading thinkers in contemporary philosophy of mind, Jaegwon Kim, observes that '[t]hrough much of the twentieth century, especially during the second half, debates over the mind-body problem were shaped by physicalism', that is, the metaphysical belief that 'all things that exist in this world are bits of matter and structures aggregated out of bits of matter'. Once this metaphysical framework is accepted, Kim goes on to say, then 'the foremost metaphysical problem about the mind is where in the physical world our minds fit – in fact, whether minds have a place in such an austerely physical world at all' (129).

Kim has his own view as to how minds can be accommodated within a physical universe. His suggestion is that the most salient features of consciousness – cognitive intentional states like belief, desire, or memory – can be coherently integrated within a physicalist worldview by means of functional reductions, that is, by re-conceptualising them in terms of the causal work they are supposed to do. To illustrate: part of the causal work of my belief that fire burns is to keep my body at a reasonable distance from it; thus, a capability to perform this function would be an essential part of any definition of what it is to believe that fire burns. Kim has little doubt that such reductions are possible in principle, even though they may be very difficult to achieve in fact. He also frankly admits, however, that there are aspects of consciousness – specifically, its qualitative, *phenomenal* aspects – that cannot be functionally re-conceptualised and are therefore physically irreducible: 'Qualia…' – he remarks – 'are the "mental residue" that cannot be accommodated within the physical domain' (143). According to Kim, the existence of this mental residue is not a serious threat to physicalism; as he has it, '[p]hysicalism is not the whole truth, but is the *truth near enough*, and *near enough* should be *good enough*' (146).

This is a striking statement. Drawing on his deep knowledge of the philosophical debate over the mind-body problem in the last decades, Kim constructs a powerful argument in support of reductive physicalism. But a theory of consciousness that fails to account for those

[1] *The Future for Philosophy*, edited by Brian Leiter (Cambridge: Cambridge University Press, 2004), 129–52.

aspects of it that we most cherish and care about (surely, it is the *painfulness* of pain that we seek to avoid!) has failed to grasp something essential to its object of investigation. One does not need to be an old fashioned Hegelian or a Bradleyan idealist to recognise that '*Das Wahre ist das Ganze*' or that 'no partial truth is really quite true': unless *all* aspects of a given phenomenon (consciousness or anything else) are accounted for, we can't be sure that we have explained it correctly. Hence, there is at present still room open for arguing that the future for the philosophy of mind cannot be just that of labouring the details of a physicalist theory of the mental, but it must also be that of inquiring further into the very nature of consciousness and the metaphysical reliability of the physicalist framework.

The papers collected in this volume provide in depth-explorations of the nature of subjectivity, the phenomenological structure of the self, its status as a metaphysical entity, its mode of interaction with, and existence within, a physical universe. As it should be, they display a variety of different perspectives and basic philosophical commitments. Taken together, they reflect not solely the systematic but also the *historical* complexity of the issue. The names of, among others, Plotinus, René Descartes, Gottfried Wilhelm Leibniz, Immanuel Kant, William James, and Edmund Husserl figure prominently in several of the following discussions as those of thinkers who still have much to contribute to the contemporary debate: as masters to be learnt from rather than, as is too common in the analytical literature, as sparring-partners who can be easily defeated.

The articles were originally presented at a conference organised at the University of Edinburgh to honour the work and philosophical legacy of Timothy L. S. Sprigge (1932–2007), former Professor of Logic and Metaphysics (1979–1989) in that same University and author of insightful works in the philosophy of mind and in speculative metaphysics. His original combination of panpsychism and absolute idealism – the theory that reality is a single whole that appears in the guise of a plurality of monad-like centres of experience – is a nice alternative, if not an antidote, to the scientistic physicalism that, as Kim observes, 'shaped' the debate over the mind-body problem in the last decades and that threatens to become the default position in contemporary philosophy of mind. Systems of revisionary metaphysics run counter to established assumptions, and rightly so; at the very least, they are an invitation and a challenge to further inquiry.

Neither the conference nor this volume would have been possible without the help of several persons and institutions. The Editors

wish to acknowledge the help received from the Royal Institute of Philosophy, the University of Edinburgh, the British Society for the History of Philosophy, the Mind Association and the Scots Philosophical Association. A special word of thanks goes to Andy Clark, Jesper Kallestrup, Leemon McHenry and Dory Scaltsas for their encouragement and constant support.

The Editors

Sprigge's Ontology of Consciousness

LEEMON McHENRY

Abstract
Timothy Sprigge advanced an original synthesis of panpsychism and absolute idealism. He argued that consciousness is an irreducible, subjective reality that is only grasped by an introspective, phenomenological approach and constructed his ontology from what is revealed in the phenomenology. In defending the unique place of metaphysics in the pursuit of truth, he claimed that scientific investigation can never discover the essence of consciousness since it can only provide descriptions of structure and function in what we normally think of as physical existence. In this paper I present a critical evaluation of Sprigge's view focusing in particular on his conception of the nature of scientific inquiry vis-à-vis the ambitious project of his metaphysics. I argue that a naturalistic metaphysics provides a more adequate approach to the relation between science and metaphysics.

1. Introduction to Sprigge's Philosophy

Timothy Sprigge arrived at a solution to the problem of consciousness that became the basis for a metaphysical system in the grand style of Spinoza, Hegel, Bradley and Whitehead. He said he came to his view by a transformation of his philosophical outlook, one that enabled him to escape from a variety of absurdities into which he had been led by premises he did not think to question.[1] By the early 1970s he had been influenced by the most respected figures in British philosophy, Bertrand Russell, G. E. Moore and A. J. Ayer, but it was the serious study of the American philosophers, George Santayana and William James, who provided a more intellectually satisfying answer to his questions, and after a long period of gestation, the impact of Spinoza and Bradley put the final touches to his system. This transformation is much in evidence from the publication of his first book *Facts, Words and Beliefs* (1970) to *Santayana* (1974) and then with the publication of his *magnum opus*, *The Vindication of Absolute Idealism* (1983). In short, Sprigge traded the relative security and respectability of mainstream analytical philosophy for what he

[1] T. L. S. Sprigge, 'Consciousness' in *The Ontological Turn*, edited by M. S. Gram and E. D. Klemke (Iowa City: Iowa University Press, 1974), 115. Also see 'Orientations', *Philosophical Writings* **2** (1996), 98–100.

doi:10.1017/S1358246110000093

knew would be widely perceived as a retreat into soft-minded, muddle-headed metaphysics.

Sprigge began his career at the University of Cambridge where he worked with R. T. H. Redpath and then A. J. Ayer at University College, London. He became Lecturer and then Reader in philosophy at the School of English and American Studies at the University of Sussex, Regis Professor of Logic and Metaphysics at the University of Edinburgh and finally Emeritus Fellow retiring in Edinburgh and then Lewes, Sussex. It was in this final period that he produced three studies of the same large scale as the Victorian novels he used to read, *James and Bradley* (1993), *The God of Metaphysics* (2006) and *The Phenomenology of Thought* (unfinished). He always preferred the sustained chain of thought in book-length monographs to journal articles and claimed he seldom read the latter in spite of the fact that he himself contributed over a hundred articles and book chapters in the course of his career. In addition to his scholarly work on Spinoza, Bentham, James, Royce, Santayana, Bradley and Whitehead, and the construction of his own original system of metaphysics he called 'panpsychistic absolute idealism', Sprigge also produced important work in utilitarian ethics and was a leading figure in the animal rights movement in Britain.

In this paper I will examine the main features and motivating factors in Sprigge's view of consciousness, explain some of the metaphysics he constructed from his initial discovery and then offer a critical evaluation of the results. His philosophy is filled with creative connections between the thinkers he absorbed, unique insights into how these positions form a consistent whole and innovative arguments for the views advanced, but as with any system of philosophy, there are problems that need to be clearly exposed and debated.[2] Enthusiasm is no homage to genius.

2. The Irreducibility of Consciousness

W. V. Quine once said of philosophical curiosities about which he was in no position to cope: 'Closer to home, there are the mysteries of the mechanism of memory; also of consciousness – I am even at a loss for a satisfactory statement of what consciousness is, though nothing is

[2] For a critical evaluation of Sprigge's thought, see P. Basile and L. McHenry (eds), *Consciousness, Reality and Value: Essays in Honour of T. L. S. Sprigge* (Frankfurt: Ontos Verlag, 2007).

more familiar'.[3] Regarding nothing more familiar, Sprigge claimed that he and 'Thomas Nagel offered, independently of each other, a definition of consciousness which amounts to saying that an individual is conscious at a certain time if there is an answer to the question what it was like being him, her, or it, at that time'.[4] In fact he had published this view three years before[5] Nagel's highly influential article, 'What Is It Like to Be a Bat?'[6] He said that this essentially non-physical view of consciousness had been obvious to him since he had been engaged in serious philosophy and believed it long before he first used it in print.[7] As Sprigge put it in a slightly more refined manner: 'An individual's state of consciousness at any moment is that concerning the character of which one is wondering if one asks oneself what it is like to be them just then'.[8] This is obviously not an epistemological matter of our attempt to know what it is like to be some individual or creature, nor is an issue of how much of one's self-consciousness one currently grasps, but rather an ontological matter concerning the reality of subjective experience. In what seems like a direct reply to Quine, Sprigge claimed that 'Whenever the question 'What is consciousness' is asked, as though consciousness were something very mysterious, one can be sure that the treatment of the subject is likely to be unsatisfactory. For consciousness is the most familiar of realities and its nature should be easier to grasp than is the nature of anything else'.[9] Should there be a lack of clarity in this formulation, Sprigge claimed one gets nearer to such a statement of the essence of consciousness as that in which being is one with its own non-discursive

[3] W. V. Quine, *Quine in Dialogue*, edited by D. Føllesdal and D. Quine (Cambridge: Harvard University Press, 2008), 28.

[4] T. L. S. Sprigge, 'The Importance of Subjectivity: An Inaugural Lecture', *Inquiry* **25** (1982), 146.

[5] T. L. S. Sprigge, 'Final Causes', *Proceedings of the Aristotelian Society* **71** (1971), 166–8.

[6] T. Nagel, 'What Is It Like to Be a Bat?', *Philosophical Review* **83** (1974), 435–50. See also Nagel's acknowledgement to Sprigge in *The View from Nowhere* (New York: Oxford University Press, 1986), 15n. Nagel also acknowledges in the same note that B. A. Farrell had asked in 1950 what it would be like to be a bat but dismissed the difficulty for materialism.

[7] T. L. S. Sprigge, 'Consciousness', *Synthese* **98** (1994), 79.

[8] T. L. S. Sprigge, 'Is Consciousness Mysterious?', *Anthropology & Philosophy* **3:2** (1999), 6.

[9] *Ibid*., 5.

knowing of itself.[10] This is sometimes called 'phenomenological consciousness'. [11] It is the view that what makes a state phenomenally conscious is that there is something 'it is like' to be in that state. For both Sprigge and Nagel, this is to assert that consciousness has an essentially inner, subjective nature that is not specifiable in any physical terms. Consciousness is available to us simply because we are conscious. We know it as something that has its own appreciative sense of itself rather than by description.

The inner subjectivity of what it is like to be a person and 'the fact that he is conscious does not follow logically – however firmly it follows inferentially – from any fact about him of a physical kind, whether it concern the patterns of his behaviour, the occurrence of various computer type processes within his brain, or whatever'.[12] The notion was reinforced in *The Phenomenology of Thought* when Sprigge recounted a point made by John Wisdom to the effect that the concept of mind has little or nothing to do with the concept of the brain.[13] After the death of some eminent thinker, let us say Albert Einstein, his brain is examined. But to everyone's amazement all that there is inside his head is hay! The very bases of science seem destroyed, but no one decides that Einstein never had any thoughts. We all believe that consciousness, in the normal case, at least, depends on the brain, however, our conception of the brain does not imply anything about consciousness and vice versa. Was the concept of consciousness enriched by the discovery of nerve-cells seen in brain slices or the discovery of the electrochemical activity of the neurotransmitters in the same way as our concept of combustion was enriched by the discovery of oxygen? Sprigge argued that it did not. While there is a causal connection between the grey and white gooey stuff normally found in a person's skull and consciousness, there is no logically necessary relation between the two, and the concept of consciousness remains relatively untouched by any advances in the neurosciences.

Sprigge's discovery about consciousness might very well seem obvious to us in retrospect, and this I suggest is because of the

[10] T. L. S. Sprigge, 'The Importance of Subjectivity', *op. cit.*, 146–7. Sprigge credits F. H. Bradley with articulating this view. See also his 'Knowledge of Subjectivity', *Theoria to Theory* **14** (1981), 320.

[11] N. Block, 'On a Confusion about a Function of Consciousness', *Behavioral and Brain Sciences* **18** (1995), 227.

[12] T. L. S. Sprigge, 'Consciousness', *op. cit.*, 79.

[13] T. L. S. Sprigge, *The Phenomenology of Thought*, edited by L. McHenry (unpublished manuscript, Sprigge Archives, Edinburgh University Library), Part 2, Chapter 1, 17.

influence that Nagel's paper has had over thirty years of discussion, but in the late 1960s when Sprigge was making his case, he said: 'a striking feature of contemporary English philosophy is the insistence that philosophers since Descartes have been led astray by incoherent notions of a realm of private experience…'.[14] Such was the stranglehold of logical positivism and behaviourism at the time. The dominant views of mind shaped by Ryle and Wittgenstein amounted to a denial of consciousness in the sense in which its existence is taken as a basic fact of the reality of our subjective experience.[15] But there is, Sprigge argued, nothing incoherent or absurd about an intrinsic quality exemplifiable by private experience. In fact, following Descartes on this one point, he agrees our knowledge of our own consciousness is basic and certain. At each moment there is a quite definite character, however easy or difficult it is to identify this character about the private aspects of one's stream of consciousness. So, there is something it is like to be a person, a cat, and indeed probably almost all of the creatures in the animal kingdom. The main difference between human consciousness and other sentient creatures is that our consciousness is largely though not exclusively preoccupied with linguistic discourse or various types of symbolic representations of our experience. But it is quite clear that there is a definite character of the experience of other creatures even if at some very low level it is merely one of a dull throb of existence, a faint awareness or a vague sense of passage. Sprigge did not think that there was something it is like to be a mere aggregate like a cloud or a mountain even though in his view there is something it is like to be one of the ultimate constituents of all things and indeed the whole universe itself which he conceived of as one single substance in the fashion of Spinoza and Bradley.

The view of experience advanced by Sprigge begins with an approach to the study of consciousness that was pioneered by William James and Edmund Husserl.[16] He thought that the phenomenological approach was the only way to understand consciousness and that no one has come closer to illuminating its nature than James. What it is like to be a conscious human being is best grasped from within and described in watery metaphors of drops of

[14] T. L. S. Sprigge, 'The Privacy of Experience', *Mind* **78** (1969), 512.

[15] T. L. S. Sprigge, 'Consciousness', *op. cit.*, 80.

[16] W. James, *The Principles of Psychology*, Vols 1 & 2 (London: Macmillan and Co, 1901); E. Husserl, *The Phenomenology of Internal Time Consciousness*, translated by J. S. Churchill (Bloomington: Indiana University Press, 1964).

experience flowing together in a stream of mental discourse, continually fading into the past and advancing into the future. The lesson from James is 'a habit of directing one's attention at a certain reality, one's own total experience, and bearing the character of what one then notes vividly in mind whenever theorizing about mind or related matters'.[17] The crucial method here involves introspection, empathy, and insight into the necessities present in what this approach reveals. As James says: 'That unsharable feeling which each one of us has … may be disparaged for its egoism, may be sneered at as unscientific, but it is the one thing that fills up the measure of our concrete actuality, and any would-be existent that should lack such a feeling, or its analogue, would be a piece of reality only half made up'.[18]

One becomes especially aware of James' accomplishment by seeing how badly philosophy is done in complete neglect of this approach, as for instance one finds in the sort of identity theory of mind and brain in which consciousness is treated as a theoretical entity invoked to explain certain sorts of behaviour. All the wrong turns in contemporary philosophy of mind that have obscured our path to the essence of consciousness have resulted from preoccupations with abstractions rather than with absorption in the lived character of conscious experience. This includes what Daniel Dennett has referred to as a 'tormented snarl of increasingly convoluted and bizarre thought experiments … and a bounty of other sidetrackers and time-wasters',[19] though obviously Dennett and Sprigge are odd bedfellows in this complaint. For Sprigge the phenomenological study of consciousness is the foundation of other inquiries. Thus unlike Nagel's limited point, Sprigge developed this crucial insight about the essence of consciousness in a systematic manner both for his metaphysics and ethics. Conscious states, or more generally sentient states, with their subjective, first-person ontology are the true reality.

So, for Sprigge, this is the answer to the not-so-mysterious question 'what is consciousness?' As for the more difficult question of how consciousness is related to the brain, he argued that the first thing to know in charting the relation is to know what each term is,

[17] T. L. S. Sprigge, 'The Distinctiveness of American Philosophy' in *Two Centuries of Philosophy in America*, edited by Peter Caws (Oxford: Blackwell, 1980), 205.

[18] W. James, *The Varieties of Religious Experience* (Harmondsworth: Penguin, 1982), 499.

[19] D. Dennett, *Consciousness Explained* (Boston: Little, Brown and Company, 1991), 369.

and that he had illuminated the nature of the former in his phenomenological consciousness. Given that the relation of mind and brain is one of identity, and given further that Sprigge has been successful in capturing the essence of consciousness as subjective inwardness, then for him the substance of brain activity must consist of particulars that have the same character.[20] From this point of departure, he constructed a panpsychist ontology in which consciousness or sentience is omnipresent in the universe.

3. The Limits of Scientific Investigation

The problem of consciousness, as Sprigge saw it, is that it refuses to cooperate with the prevailing materialistic paradigm; it is the stubborn irreducibility of consciousness or the failure to explain how neurophysiological processes of the brain have any connection at all to that nothing more familiar phenomena of consciousness that remains the major stumbling block. It is, he said, the problem that has 'bedeviled philosophy'.[21] More recent writers have identified this as the 'hard problem', one that might very well be intractable because human beings simply don't have the intelligence or cognitive tools designed for understanding consciousness in the same way dogs don't have enough wattage in the bulb to do mathematics.[22] At least here we seem to know enough to know what we can't know, but Sprigge disagreed. Somehow this attitude of the 'new mysterians' such as Nagel and McGinn has escaped notice that it is defeatist. Why aren't all the unending philosophical problems just like the hard problem of consciousness equally intractable or hopeless? In Sprigge's idealist reversal of the hard problem, it is not consciousness that we fail to understand; rather in a move reminiscent of Berkeley, it is the very concept of physical matter that we do not understand. So, any notion that the nature or essence of consciousness is something that awaits the results of neurophysiological research to be understood is fundamentally misguided.

Sprigge argued that our knowledge of our own states of consciousness which are irreducibly subjective and private will be threatening

[20] T. L. S. Sprigge, 'Is Consciousness Mysterious?', *op. cit.*, 19.

[21] T. L. S. Sprigge, *The Phenomenology of Thought*, *op. cit.*, Part 1, Chapter 1.

[22] C. McGinn, *The Mysterious Flame: Conscious Minds in a Material World* (New York: Basic Books, 1999), 47.

only to those absorbed in current physicalistic dogma because such states present them with the possibility of failure to give a complete account of knowledge about the world. Some physicalists have even tried to convince themselves that they are not, in fact, conscious. The very title of Sprigge's major work, *The Vindication of Absolute Idealism*, was a challenge to such philosophers for he thought that physicalism had become such a dogma that it had shut down discussion of other plausible options even if they were thought to be buried once and for all by the critiques of Russell, Moore and Ayer. The fact that we can know things about our own states that others cannot know is knowledge of truth about the world that even the most complete description of how things are in reality would miss out. This is the essence of the argument from phenomenological consciousness. Given all the physical information about a person, a bat or whatever, we still would not know what it was like to be that person or creature. Something would be left out, but it is not something physical, hence physicalism, the thesis that all reality must be physical reality, must be false.

Scientific investigation can no more reveal the essence of consciousness than the congenitally deaf can hear music or the congenitally blind can see colours. Our knowledge of the physical world remains nothing but structure and function no matter how far down or up we go with our microscopes and telescopes. Recent developments in neuroimaging reveal much fascinating information about the brain and might possibly aid in our understanding of affective disorders, but still reveal nothing about what it feels like to suffer from depression. Computer programmes might tell us much about cognitive functioning, but we still remain in the dark about how the brain thinks or what it is like to be a thinking thing. Sprigge thus armed with his distinction between the World of Description and the World of Acquaintance argued that scientific investigation is limited to investigations of appearances rather than the underlying reality of noumenal existence.[23] Here his view has been echoed in a powerful way by David Chalmers' argument against reductive explanation. As Chalmers puts it, '…the explanation of consciousness is not just a matter of explaining structure and function. Once we have explained all the physical structure in the vicinity of the brain, and we have explained how all the various brain functions are

[23] T. L. S. Sprigge, 'The World of Description and the World of Acquaintance' in *Beyond Conflict and Reduction: Between Philosophy, Science and Religion*, edited by W. Desmond, J. Steffen and K. Decoster (Leuven: Leuven University Press, 2001), 23.

performed, there is a further sort of explanandum: consciousness itself'.[24] Other philosophers such as Russell have also made this point about the limits of scientific investigation and the non-inferential nature of our knowledge of the mental. 'Physics', he said, 'is mathematical not because we know so much about the physical world but because we know so little; it is only its mathematical properties that we can discover. For the rest, our knowledge is negative…'.[25] But Sprigge's view is not simply that science fails to get at the inner essence of what is perceived as physical reality. It is rather that science and metaphysics have two completely different aims. His view in this respect parallels Bradley who espoused a strong distinction between appearance and reality. Empirical science is confined to the former while metaphysics attempts to uncover the true or absolute nature of the latter. In other words, science is primarily concerned with abstractions whereas metaphysics is focused on revealing the nature of concrete reality.[26] What Sprigge meant by 'metaphysics' however is clearly not just any metaphysics but rather the sort that begins with affirming the subjective reality of consciousness. Accordingly, he argued that the World of Acquaintance is more real than the World of Description. Everything that exists must have its own appreciative sense of its own being and of its relation to others. What actually presents itself to us as the physical world is really a system of just such self-and-other-appreciating individuals whose togetherness forms one eternal cosmic consciousness.[27]

4. The Argument for Panpsychism

In his *Vindication of Absolute Idealism*, Sprigge argued that every conception of a physical reality is on deeper reflection really a phenomenological conception of reality, and corresponding to every phenomenal appearance there must be a noumenal backing that is

[24] D. Chalmers, *The Conscious Mind* (New York: Oxford University Press, 1996), 107.

[25] B. Russell, *An Outline of Philosophy* (London: Routledge, 1992), 125. Also see *Human Knowledge: Its Scope and Limits* (London: George Allen and Unwin, 1948), 240.

[26] F. H. Bradley, *Appearance and Reality* (Oxford: Clarendon Press, 1893), 250–51 and T.L.S. Sprigge, *The God of Metaphysics* (Oxford: Clarendon Press, 2006), 476–8.

[27] T. L. S. Sprigge, *The Vindication of Absolute Idealism* (Edinburgh: Edinburgh University Press, 1983), Chapter 6 and *The God of Metaphysics*, *op. cit.*, 486–90.

essentially consciousness.[28] Whereas Kant had claimed that the noumenon or the thing-in-itself is essentially unknowable since it lies outside the categories of the understanding and sensible intuition,[29] Sprigge developed the idea that whatever we understand by physical reality, it must have a noumenal backing or being that is essentially psychical in character. The sort of knowledge one has in knowing oneself as a centre of consciousness reveals the inherent nature of one concrete reality or one such noumenon behind the phenomenon.[30] Once one has a sense of what Sprigge meant by the noumenal reality of human consciousness, it is then relatively easy to understand how he extended this concept to other beings as the noumena behind the phenomena perceived as physical reality. This he articulated as an ontology of distinct units, momentary centres of experience, that make up nature but are also related in such a way as to compose wholes which may or may not be sentient or conscious themselves.

Sprigge argued first in a Berkeleyan fashion that there is no such thing as an unexperienced physical reality because none of us, he claimed, can conceive of a world divested of consciousness. As he said: 'Whenever we seriously try to bring into clear consciousness the nature of any situation in the existence of which we believe, we imagine it either as a form or a content of some consciousness, actual or possible.'[31] But nature does not exist only as an object for human consciousness or as a system of possible sensations on our part. This leaves him with the view that for something to exist it must be experience.

Put in its simplest form, the crucial argument for this position is based on two premises:

1. That the physical world, or at least something which is the reality which we conceive as the physical world, exists independently of the consciousness of it by humans and animals, and that it existed before there were any humans or animals.
2. That nothing can exist except sentient experience.

From which it follows that:

3. The physical world is composed of sentient experience.[32]

[28] T. L. S. Sprigge, *The Vindication of Absolute Idealism, op. cit.*, 39.

[29] I. Kant, *Critique of Pure Reason*, translated by N. Kemp Smith, (London: Macmillan & Co., 1963), A 236/B 295–A 260/B 315.

[30] T. L. S. Sprigge, *The Vindication of Absolute Idealism, op. cit.*, 40–1, 105–10.

[31] T. L. S. Sprigge, 'Are There Intrinsic Values in Nature?', *Journal of Applied Philosophy* **4** (1987), 25.

[32] T. L. S. Sprigge, *The God of Metaphysics, op. cit.*, 484.

The first premise distinguishes Sprigge's view from Berkeley's subjective idealism and commits him to a form of realism. For Berkeley there are two sorts of things: minds (divine and human) and ideas whereas for Sprigge there are innumerable streams of experience existing independently of human and animal consciousness. But as is implied in the second premise what is external to us is nothing but sentient experience. This is a bold claim since it appears he neglects the unlikelihood of any trace of consciousness or sentience in the explosive beginning of the universe in the Big Bang or the magma oozing from a volcano. There is also the question about whether consciousness or sentient experience needs a brain. Panpsychism has thus been frequently characterized as lovely and tempting but less than sober.[33] Yet according to Sprigge it provides the best explanation for the answer to the question of the noumenal backing of phenomena and a solution to the mind-body problem. What is required in understanding the second premise is an obvious extension of the term 'experience' such that it is not understood in terms of human perception. Royce seems to come closest to making sense of this idea with his notion of apperceptive time spans.[34] What we perceive as inorganic nature is really a nature alive to various degrees. Another is Whitehead's concept of causal efficacy as the basic mode of perception in nature. As he argued:

> A jellyfish advances and withdraws, and in so doing, exhibits some perception of causal relationships with the world beyond itself; a plant grows downwards to the damp earth, and upward towards the light. There is thus some direct reason for attributing dim, slow feelings of causal nexus, although we have no reason for any ascription of the definite percepts in the mode of presentational immediacy….
>
> As we pass to the inorganic world, causation never for a moment seems to lose its grip. What is lost is originativeness, and any evidence of immediate absorption in the present.[35]

What these philosophers sought to achieve was a radical revision of classical British empiricism, one that naturalized mind and mentalized nature. Sprigge followed very closely these views when he formulated his ontology of momentary centres of experience.

[33] See C. McGinn, 'Hard Questions' in *Consciousness and its Place in Nature*, edited by A. Freeman (Exeter: Imprint Academic, 2006), 93.

[34] J. Royce, *The World and the Individual*, second series (New York: Dover, 1959), 226–7.

[35] A. N. Whitehead, *Process and Reality*, *op. cit.*, 249.

5. Recent Developments

Sprigge divided philosophical positions into two classes: (1) *incredible* because they contradict basic premises regarded as personal certainties by those who hold them, and (2) *serious options* because they supply answers to questions, the holding of which does not require an extraordinary leap of imagination.[36] In his own time Sprigge found himself among philosophers who regarded his position as incredible for his holding the view that consciousness is the basic certainty, and more specifically the existence of consciousness as something not explicable in physical terms. We have come a long way in that the number of philosophers who no longer regard his views as incredible has increased significantly. Given the stalemate in explaining consciousness in contemporary philosophy of mind, even panpsychism has been re-introduced into mainstream by the likes of David Chalmers, Galen Strawson, David Skrbina, and William Seager.[37] Galen Strawson has breathed new life into panpsychism with his argument that physicalism, rightly understood, entails panpsychism. He begins like Sprigge with a realism about experience, consciousness or what-it-is-likeness that commits him to the idea that the experiential is physical and thereby challenges Descartes' notion of the unextended mind.[38] David Chalmers has flirted with the notion by suggesting that any physical process can be treated as an information-processing system which is realized both physically and phenomenally.[39] There is even some question here as to whether Andy Clark's extended cognition thesis leads to extended consciousness.[40] For Sprigge the skull and skin are no boundary since our experience is just a fragment, a stereoscopic perspective of a field, which must be conceived as sentience lying behind the extended four-dimensional space-time. When physicalist philosophy

[36] T. L. S. Sprigge, 'Consciousness', *op. cit.*, 73.

[37] See for example, D. Chalmers, *The Conscious Mind*, *op. cit.*, 276–310; G. Strawson, 'Realistic Monism: Why Physicalism entails Panpsychism' in A. Freeman (ed.), *Consciousness and its Place in Nature, op. cit.*, 3–31; D. Skrbina (ed.), *Mind that Abides: Panpsychism in the New Millennium* (Amsterdam: John Benjamins, 2009); W. Seager, 'Consciousness, Information, and Panpsychism', *Journal of Consciousness Studies* **2:3** (1995), 272–88.

[38] G. Strawson, 'Realistic Monism', *op. cit.*, 3–31.

[39] D. Chalmers, *The Conscious Mind, op. cit.*, 288–99.

[40] This is a view that was considered but rejected in A. Clark and D. Chalmers, 'The Extended Mind', *Analysis* **58** (1998), 7–19, yet the authors agree the extended mind implies an extended self.

of mind begins to make overtures to panpsychism we know we have entered a new period of discourse. Even if its new adherents embrace the doctrine with a certain amount of reluctance and caution, it seems they no longer feel they need to disguise their views when appearing among their orthodox analytical colleagues. Sprigge never needed to do so; he was unabashed in his commitment to panpsychism.

6. Critical Evaluation

Beginning with the irreducibility thesis and ending with a full-blown panpsychistic absolute idealism has a certain respectable coherence and explanatory power. It solves certain problems and avoids others, but then creates obvious problems of its own, many of which are generic to idealism. While I think Sprigge's argument about phenomenological consciousness is as good as one finds in philosophy of mind, I have always found his panpsychism fascinating yet fraught with difficulties. Aside from presenting a strong case against physicalism, it is less clear that he provided convincing justification for the second premise of his argument for panpsychism, *i.e.* that nothing can exist except sentient experience. That is, granted the Berkeleyan argument that to be is to be perceived – *esse est percipi* – or that objects of experience exist only in perceivers, how do we then get all objects of nature are themselves subjects of experience? Second, in spite of the attempts by Royce, Whitehead and Sprigge, it is very hard to understand how the use of the terms 'experience', 'sentience' and 'consciousness' can be meaningfully applied to inanimate nature. Our powers of imagination or empathy fade once we are outside of the realm of animal sentience. And finally, there are still far too many explanatory gaps to understand the emergence of the high grade streams of consciousness from the low grade momentary centres of experience even though the theory has the overall advantage of providing a smooth and continuous interpretation of nature in term of one ontological type. Sprigge thought we needed a panpsychist philosophy of science, but admitted he was not prepared to do the job. But in what follows I have a broader concern about his view of the relation, or lack of relation, between science and metaphysics.

Regarding Sprigge's view of the aims of science, I suspect many scientists would find it unsatisfactory to describe their objectives as restricted to mere appearances, and that a special group of thinkers called metaphysicians have a privileged access to reality that reveals the truth about the world. Aside from remnants of positivism

surviving in current science it is hard to see how science is confined to this severely limited role. Scientific theories often pass beyond phenomena altogether to postulate an unseen world whose detailed workings explain and unite diverse phenomenal effects. The quest for a Theory of Everything in theoretical physics incorporating such speculative theories as quantum ontologies, quantum theories of gravity, superstring theory and the multiverse hypothesis are prime examples. Theories of speculative biology seek explanations of the origin of life beyond mere structure and function. So, while it may seem useful to make a sharp distinction between appearance and reality, it is implausible that science and metaphysics operate in two realms, serve different purposes or have different functions.

As opposed to overly optimistic claims of pure metaphysicians a more defensible role for the discipline has been articulated by naturalized metaphysicians who see such theorising as merely the general end of theory.[41] Metaphysics is therefore continuous with science and risks its dignity in the same way. As Quine made the point: '[Naturalism] sees natural science as an inquiry into reality, fallible and corrigible but not answerable to any supra-scientific tribunal, and not in need of any justification beyond observation and the hypothetico-deductive method'.[42] Naturalized metaphysics rejects the idea of some other way to truth independent of science, but this is not necessarily an affirmation of physicalism, since the specific ontology will be an open question. It might be panpsychist if, for example, the philosopher or scientist sees it as the best, tentative solution to the problem of emergence or the mind-body problem. It might be monistic in the fashion of Einstein, pluralistic as in the case of Heisenberg, an Aristotelian substance ontology or a Whiteheadian event ontology. Where it seems to me, however, that Sprigge was right is in thinking that scientific investigation has nothing to contribute to the understanding of the intrinsic nature

[41] This distinction appears in my *Whitehead and Bradley: A Comparative Analysis* (Albany: State University of New York Press, 1992), 16–9. Pure metaphysicians would include Bradley and Sprigge who argue for a sharp separation of science and metaphysics whereas naturalized metaphysicians would include Peirce, Whitehead and Quine who see metaphysics and science as necessary to rounding out our general theory of the world.

[42] W. V. Quine, *Theories and Things* (Cambridge, Mass: Harvard University Press, 1981), 72. Quine espoused a thoroughgoing physicalism and behaviourism in line with his naturalism. Private, subjective phenomena play no role in his ontology, which perhaps explains why he regarded consciousness as a mystery.

of conscious experience, but this has more to do with the limits of empirical investigation and the recognition of the role of introspective psychology than with metaphysics *per se*. Sprigge himself much admired scientific metaphysics but was also concerned to demonstrate that metaphysics could arrive at truth independent of science. It is this assertion that I find most problematic in his system. That is, it seems to me one claim that a phenomenological approach best grasps the nature of consciousness and quite another that metaphysical speculations from this point of departure reveal truth about the universe. Once experience has been interpreted in terms of ontological categories, one has left the relatively safe haven of phenomenology and entered the realm of metaphysical speculation. It is this leap between the what-it-is-likeness of consciousness to a panpsychist ontology that begins the process of theorising, one in which I find science and metaphysics to be engaged in the same quest. Sprigge equated a phenomenological description or state with a metaphysical truth without recognising the heavy dose of ontologising involved in the interpretation of experience. For example, this becomes clear in his very use of the concept of the noumenon to describe the subjectivity of experience which then becomes the basis for the ontological units of his system, the momentary centres of experience. Not that I deny him the ontological interpretation of experience or the leap from the phenomenological to the ontological, but however plausible or implausible it is, it needs to be recognized as such.

The view of science as mere abstractions may strike us as an obvious defect of idealism that warranted a strong revival of materialistic philosophy of mind. It was after all the failure of idealism to serve as an adequate foundation for the sciences that led to its demise at the beginning of the twentieth century, but Sprigge need not propound the strong separation of metaphysics and science to make his case for how much science is limited to revealing the essence of consciousness.

7. Conclusion

To engage in the project of systematic metaphysics at the time in which Sprigge did was a bold endeavour and one that ran so much against the grain of respectable philosophy that he faced isolation and ridicule. He demonstrated, however, through the sheer reasonableness of his arguments the strength of his view of consciousness which raised serious doubts about the complete success of the physicalist programme. Explaining consciousness is no mystery as long as one begins with one's own phenomenology rather than the theoretical

constructs of physicalism, functionalism or behaviourism. Commitment to phenomenological consciousness, however, does not, in my view, privilege metaphysics against science as the only discipline with access to truth. [43]

California State University, Northridge

[43] I have profited from discussions with Frank McGuinness, Pauline Phemister and Pierfrancesco Basile on earlier drafts of this paper.

How to Become Unconscious

STEPHEN R. L. CLARK

Abstract
Consistent materialists are almost bound to suggest that 'conscious experience', if it exists at all, is no more than epiphenomenal. A correct understanding of the real requires that everything we do and say is no more than a product of whatever processes are best described by physics, without any privileged place, person, time or scale of action. Consciousness is a myth, or at least a figment. Plotinus was no materialist: for him, it is Soul and Intellect that are more real than the phenomena we misdescribe as material. Nor does he suppose that consciousness depends on language (as Stoics and modern materialists have sometimes said): wordless experience is actually superior. And much of what counts towards our present consciousness is to be discarded. It is better *not* to remember most of what now seems more significant to us; better to discard images; better that the intellect be 'drunk' than 'sober', losing any sense of separation between subject and object. The goal of the Plotinian intellectual is to join 'the dance of immortal love', but it is a mark of the good dancer that she is not conscious of what she does. There is therefore a strange confluence between Plotinus and modern materialists: *our* experience at least is transitory, deceitful, epiphenomenal, and 'reality' is to be encountered when we have shed our illusions.

1. The Materialist Problem

Consistent materialists are almost bound to suggest that 'conscious experience', if it exists at all, is no more than epiphenomenal. A correct understanding of the real requires that everything we do and say is no more than a product of whatever processes are best described by physics, without any privileged place, person, time or scale of action. Consciousness is a myth, or at least a figment, and the real causes even of what we seem to ourselves to think and do lie out of sight and hearing, in the motions and chemical reactions of our brains. Really consistent materialists are prepared to say (if 'saying' is anything that they actually believe they do) that 'conscious selves' and choice are as fictitious as planetary angels. Less consistent, but marginally saner, materialists speak reverently of the way that new properties 'emerge' from the conjunction and association of lesser parts – though they cannot themselves have any causal impact (it's just convenient sometimes to pretend they do).[1]

[1] An 'emergent' property of a whole differs from a 'resultant' in that whereas the latter has an intelligible, mathematical relationship with the

doi:10.1017/S1358246110000056

Plotinus was no materialist: for him, it is Soul and Intellect that are more real than the phenomena we misdescribe as material, and the road to reality is through the 'inward turn', whereby we recognize the power and dignity of our souls, and realize that it is what is known through intellect that is real, rather than the flux of ego-centered sensation. It is impossible – he supposes – that either soul or intellect could 'emerge', without ever having previously existed, from the collocation of material elements. I shall not readdress either his or other arguments against the mere materialist.[2] But there is a strange convergence between Plotinian and modern materialist: the undivided origin, the One, cannot be conscious of itself, and most of what passes for our consciousness should be discarded. The fluctuating, partial We (for he usually employs the plural *Hemeis* where we might speak of the Ego[3]) is not our identity, even on the bodily plane. 'There are a great many things which the whole being does not notice.... And this is all the more so when the body is of great size, as they say happens with great sea-beasts, with whom, when a part is affected, no perception of it reaches the whole because of the smallness of the disturbance.'[4] Most of the world, including our own bodies, is managed by 'Nature', without any help from us, and we would do better not to get involved with it, or at least not to tie our fortunes to it. Nor are we necessarily aware of our 'higher self'[5] – which is no odder than remembering that we

---

properties of the parts, the former does not. There are no really convincing instances of such 'emergence' that are not simply versions of the very puzzle, the so-called Hard Problem of Consciousness, that we face: a problem created, as I suppose, by first imagining a merely material world and then being surprised that once purged of any so-called 'secondary' qualities it has no intelligible connection with those same 'secondary' (but for us immediate) properties themselves.

[2] See my 'Plotinus: Body and Mind' in *Cambridge Companion to Plotinus*, edited by Lloyd Gerson (Cambridge University Press: Cambridge, 1996), 275–91; and 'A Plotinian Account of Intellect', *American Catholic Philosophical Quarterly* **71** (1997), 421–32.

[3] See Pauliina Remes, *Plotinus on Self: The Philosophy of the 'We'* (Cambridge: Cambridge University Press, 2008).

[4] Plotinus, *Ennead* IV.9 [8].2,14–9. All translations, unless otherwise identified, are from A. H. Armstrong's edition for the Loeb Classical Library: *Plotinus: the Enneads* (London: Heinemann, 1966–88).

[5] Cf. Andrew Smith, 'Unconsciousness and Quasi-consciousness in Plotinus', *Phronesis* **23** (1978), 292–301.

are not actually aware of our bodily self much of the time, nor fully at any time, even though – for the modern materialist – that *bodily* self is what makes us what we are. In becoming Intellect, or even rising to the 'higher' level of soul, we forget or put aside all merely parochial or particular memories. Heracles' shadow, maybe, might recall his earthly life, but Heracles himself no longer minds such things.[6] The souls of the stars need not remember where they've been.[7] Plato's myth of Er, in *The Republic*, imagined that inferior souls drink their fill of Lethe, river of Forgetfulness, while superior souls, aimed at enlightenment, restrain their thirst – and remember. Plotinus, in effect, rewrites the story in a way that Dante might approve.[8] Purification is a waking up from inappropriate images,[9] and going naked into the shrine.[10] Oblivion allows us to awaken to the higher realm.[11] Even the universal Intellect itself, which is intelligible reality, exists almost in error: 'beginning as one it did not stay as it began, but, without noticing it, became many, as if heavy, and unrolled itself because it wanted to possess everything – how much better it would have been for it not to want this, for it became the second!'[12] And soul in its turn, when it comes out of the intelligible world, 'cannot endure unity, but embraces its own individuality and wants to be different, and so to speak puts its head outside, it

[6] *Ennead* IV.3 [27].27, after Homer, *Odyssey* 11.601ff. See also IV.3 [27].32, 24f.

[7] *Ennead* IV.4 [28].8, 41ff.

[8] *Purgatory* §33, 91–9; 127–9: Dante at the top of Purgatory peak drinks first of Lethe, banishing all memory of sin, and then of Eunoë, restoring an objective, guiltless knowledge of what has been.

[9] *Ennead* III.6.5, 23ff.

[10] See my 'Going Naked into the Shrine: Herbert, Plotinus and the Constructive Metaphor' in D. Hedley and S. Hutton (eds), *Platonism at the Origins of Modernity* (Dordrecht: Springer, 2008), 45–61.

[11] See Marcel Detienne, *The Masters of Truth*, translated by Janet Lloyd (New York: Urzone, Inc., 1996), 181n107, citing IV.3.32, IV.4.1 after René Schaerer, *Le Héros, sage et l'événement* (Paris: Aubier, 1965), 193–4 and Edward Warren, 'Memory in Plotinus', *Classical Quarterly* **15** (1965), 252–60.

[12] Plotinus, *Ennead* III.8 [30].8,32–36. Armstrong interprets the 'heaviness' as that of drunken sleep (as is possible), but there may be reason to suspect instead that it is the 'heaviness' of pregnancy, as Stephen Mackenna preferred in his translation of the *Enneads* (*The Enneads* edited by J. Dillon (Harmondsworth: Penguin, 1991); an abridged edition of the original Faber 1956 edition); see also *Ennead* VI.7 [38].15.

thereupon acquires memory.'[13] The things we do not know matter more for what we are than those we think we do: whatever we are, we are not *Cartesian* selves.[14] Our journey back, to Reality and Beyond, can be described as a progressive stripping away – first of words, then of memories, and even of intelligence (as that is ordinarily understood). What will be left at each stage may, in a way, be conscious – but it could as easily be described as being *unconscious*, on the far side of division. So how are we to achieve it (whatever quite it is)? And why?

2. Wordlessness[15]

First, let us consider the notion of wordless experience. One way that modern materialists have sought to explain our consciousness away is to equate it with our habit of talking about things (including what we think of as ourselves).[16] It is not a new idea. It has been a philosophical trope at least since the early Stoics that thought requires language. Those who cannot speak, it has been said, can't think. Nor can they experience or feel.[17] 'Only a creature which has a theory of mind (or, at least, a concept of experience, or of types of experience) is capable of enjoying conscious experiences, or of having phenomenal

[13] *Ennead* IV.4 [28].3: 'memory', in this context, indicates that our soul is now conscious of something that is not *present* to it in the way that reality is *present* to an eternal view (on which, more below).

[14] See Richard E. Aquila, 'On Plotinus and the "Togetherness" of Consciousness', *Journal of the History of Philosophy* **30** (1992), 7–32.

[15] Some of the material in this section was presented under the title 'Plotinus and the Sounds of Silence' at a conference on Naturalism at Bristol University, May 2002.

[16] 'Talking', they suppose (?), is simply voicing strings of symbols, rather than any strongly intentional act: once our computers can simulate a conversation 'well enough' they will be held to be 'conscious' in the only way that anything can 'really' be. Cf. Alastair Hannay, 'Comments on Honderich, Sprigge, Dreyfus and Rubin, and Elster' in *Synthese* **98** (1994), 95–112: 'If I should ever deny the existence of consciousness or begin to agree that the presently extant theories, including all extant versions of functionalism, provide adequate accounts of all we ever need to know or imply by saying we are conscious, I trust there will be some independent evidence that I have lost my mind' (99).

[17] So also in the *Hermetic Corpus* 9.1–2: 'there is no evidence of understanding without reasoned speech. … and without understanding it is impossible to have sensation' (translated by Brian P. Copenhaver, *Hermetica* (Cambridge: Cambridge University Press, 1992), 27).

feels'.[18] Most people find the conclusion grossly implausible: even if they can agree that dogs don't meditate on metaphysics, they rarely concede that dogs don't think or feel. Even Carruthers, at the close of his weirdly tendentious argument that there is really *nothing* it is like to be a bat, or cat, or horse,[19] hurriedly concedes that human infants do discriminate between experiences – and so have them – because they belong to a species which has evolved a capacity to do so (though they themselves have no theory of mind, no language and no concept of experience). Even Carruthers does not *explicitly* draw the conclusion that it is a waste of time to trouble with anaesthetising animals, and claims that it is at least wrong to do what you *think* hurts them (those who don't think anything can hurt them, presumably, can be as callous as they please).

The persistence of this argument – or so it seems to me – owes more to human egoism and a dislike of change than any rational consideration. If no-one can identify any event or object till they have a word for it (and follow the public rules for using just that word), how can they possibly acquire the knowledge? What must this be but a miracle, phylogenetically as well as ontogenetically? 'The discussions that assume that language determines thought carry on only by a collective suspension of disbelief. A dog, Bertrand Russell notes, may not be able to tell you that its parents were honest though poor, but can anyone really conclude from this that the dog is *unconscious*?'[20] Apparently, some people can. It is some comfort to me that any experimentalist who acted on Carruthers' theory (at any rate within the jurisdiction of the Animal (Scientific Procedures) Act 1986) would certainly lose her licence (and probably her job).

But what did Plotinus have to say about the possibility of wordless experience? He would agree that *human* animals have access to their own higher selves, and to the Intellect – which is to say, an intellectual grasp of the whole world – but not because they speak. The star-gods,

[18] Peter Carruthers, *Language, Thought and Consciousness: an Essay in Philosophical Psychology* (Cambridge: Cambridge University Press, 1996), 158. Carruthers goes on to suggest that conscious awareness of its own states could only evolve and persist in conjunction with a capacity to distinguish appearance and reality. So an omniscient being, presumably, for which there is no such distinction, must be unconscious.

[19] Cf. T. L. S. Sprigge, 'Final Causes', *Aristotelian Society Supplementary Volume* **45** (1971), 149–70, introducing the notion 'what it is like to be something'. See also A. L. Washburn, *The Animal Mind*, 2nd ed. (New York: Macmillan, 1917), 3ff., on what it is like to be a wasp.

[20] Steven Pinker, *The Language Instinct: the New Science of Language and the Mind* (London: Allen Lane, 1994), 58.

Plotinus is ready to suppose, live in the light of that singular cosmos, needing neither memory nor verbal speech. Non-human animals, animated by souls the equal of our own, involve those souls more deeply in the corporeal and particular, so that they have no access to the intellects whose souls they are – but not because they *don't* speak. *Human* animals, that is, live in between two sorts of entity that, paradoxically, resemble each other more than either resemble ordinary humanity. Neither have any need for human speech. Neither need have any detailed, individual memories (stars less so than animals). Neither act self-consciously, but both have experiences. Or at any rate the animals do, in that they form *images*, and respond not only to what we – but this is more contestable than we like to think – regard as 'objectively real', but also to their own images of those things. It is through *images* also that we, initially, are led to a greater understanding, though these too must be abandoned in the end.

The most obvious context for wordless experience is the mystical appreciation of the One, to which I shall return. That presence superior to being and knowledge, which sheds its light on the intelligible and phenomenal world, is of course vital to an understanding of Plotinus' system whether it is given a metaphysical or psychotropic interpretation – but it is not my only topic in this paper. For it is not only the One that lies beyond language: the Intellect itself, even in its triune mode (constituted as thinker, object thought, and thinking),[21] is wordless.

> There is no deceit There; or where could it find any thing truer than the truth? What it speaks, then, is that, and it speaks it afterwards, and speaks it in silence.[22]

> Often I have woke up out of the body to my self, and have entered into myself, going out from all other things; I have seen a beauty wonderfully great and felt assurance that then most of all I belonged to the better part; I have actually lived the best life and come to identity with the divine; and set firm in it I have come to that supreme actuality, setting myself above all else in

[21] *Ennead* V.3 [49].5, 44. All Neo-Platonists distinguish the One and the Intellect: the latter is essentially complex in that thinking and the objects thought are distinct, and the objects of thought, the whole interconnected system of intelligible reality, are many. They are unified through the presence of 'the One', which lies beyond both intellect and being (as Plato said: *Republic* 6.509b).

[22] *Ennead* VI.7 [38].34.

> the realm of Intellect. Then after that rest in the divine, when I have come down from Intellect to discursive reasoning (*logismos*), I am puzzled how I ever came down, and how my soul has come to be in the body, when it is what it has shown itself to be by itself, when it is in the body. Heraclitus, who urges us to investigate this, positing 'necessary changes' from opposite to opposite, and saying 'way up and way down' and 'changing it is at rest', and 'weariness to toil at and be subjected to the same things', has left us guessing, since he has neglected to make clear to us what he is saying, perhaps because we ought to seek by ourselves, as he himself found and sought. And Empedocles, when he said that it is a law that sinful souls should fall into this world, and that he himself has come here as 'an exile from the country of the gods' who 'put his trust in raving strife', revealed just as much as the riddling statements of Pythagoras and his followers about this, and many other matters (and besides, he is unclear because he writes poetry).[23]

Students of Plotinus may feel that he has a nerve to complain of Heraclitus' deliberate obscurity or Empedocles' poetic riddles – though the difficulty of *Plotinus'* text is almost always the effect of his attempting to explain and argue[24] – but it is at least clear that *this* 'mystical experience' is not of the One, but of the Intellect: seeing the eternal beauty of things strikes him speechless. How can it ever have happened – if we were once in full and perfect communion with the Being and Beauty of all things – that we came down into individualized existence, each with a single, distorted perspective on the world? It is as if we somehow 'got bored with being together' and each wanted our own little world more than the view, as it were, from nowhere and from everywhere.[25] Even without resort to metaphysics, we can wonder how it is that once we have seen things straight – objectively (as we now suppose) – we forget so quickly, and revert to our usual egocentric, linear and temporal experience. The momentary experience of things as they are 'in themselves' and not merely as they are 'for us' is wordless.

This claim runs counter to Richard Sorabji's reply to Anthony Lloyd on the subject of 'non-propositional thought'. According to

[23] *Ennead* IV.8 [6].1.

[24] Actually, Plotinus takes considerable pains to be clear, and criticises those who don't: see especially II.1 [40].1, II.9 [33].15, and Porphyry, *Life of Plotinus* 13, on his response to Thaumasius' wish for general statements. But his readers don't always do him justice.

[25] *Ennead* IV.8 [6].4.

Sorabji 'Lloyd ascribes to non-discursive thought four attributes which [Sorabji] think[s] belong only to the higher level of union with the One. He mentions that non-discursive thought involves no complexity, and hence (secondly) is not directed at propositions, since they are complex. He believes, thirdly, that it involves no self-consciousness, and fourthly that it is typically described in terms of contact.'[26] Sorabji argues contrariwise that noetic thought, however unifying and unified, is still of many things and hence *could* be propositional, and is also self-aware. In this, he is, I think, simply mistaken. But neither Lloyd nor Sorabji have Plotinus' reasons clear.

First, there is indeed a firm distinction between discursive and noetic thought. The soul reasons step by step towards a conclusion. The intellect grasps everything together: the distinction is like that between someone who laboriously counts how many sheep there are on the hillside, and one who simply sees.[27] Of course it is still, in Plotinus' terms, a *complex* entity: what it grasps, which is intelligible reality, is itself complex, and there is a distinction in thought between the intellect and what it thinks. That distinction *in thought* does not, of course, permit the distinct parts any independent existence: *nous, noesis* and *noeton* are all one.[28] Sorabji is right to say that non-discursive thought is complex, but quite wrong – in Plotinus' terms – to suppose it is propositional.

> One must not then suppose that the gods and the 'exceedingly blessed spectators' in the higher world contemplate propositions (*axiomata*), but all the Forms we speak about are beautiful images in that world, of the kind which someone imagined to exist in the soul of the wise man, images not painted but real. This is why the ancients said that the Ideas were realities and substances. The wise men of Egypt, I think, also understood this, either by scientific (*akribes*) or innate (*sumphute)* knowledge, and when they wished to signify something wisely, did not use the forms of letters which follow the order of words and propositions (*logoi* and *protaseis*) and imitate sounds and the enunciations of philosophical statements (*prophoras axiomaton*), but by

[26] Richard Sorabji, 'Myths about non-propositional Thought' in *Language and Logos. Studies in Ancient Greek Philosophy, presented to G. E. L. Owen*, edited by M. Nussbaum and M. Schofield (Cambridge: Cambridge University Press, 1982), 295–314: 310.

[27] Few of us can see many more than seven or eight such sheep as a *Gestalt*: angels, presumably, see any number of things like that.

[28] *Ennead* V.3 [49].5, 44. This is the real Plotinian Trinity: see also *Enneads* VI.8 [39].15, where lover, beloved and love are also unified.

> drawing images and inscribing in their temples one particular image of each particular thing, they manifested the non-discursiveness of the intelligible world, that is, that every image is a kind of knowledge and wisdom and is a subject of statements, all together in one, and not discourse (*dianoesis*) or deliberation (*bouleusis*).[29]

The intellect does not apprehend propositions *about* reality, but the realities themselves: 'there the truth is not correspondence with something else, but really belongs to each individual thing of which it is the truth'.[30] Any attempt to represent or evoke that apprehension must in its turn provide us with living (that is, self-moving) images, not words. 'Even though the soul is always moved to intelligent activity it is when it comes to be in the image-making power that we apprehend it .... [W]e are always intellectually active but do not always apprehend our activity'.[31] Words, propositions, theorems and the like are the province of a lower mode. And the images we employ must sometimes and somehow be coincident with the images that nature herself evokes, or with the realities she mirrors.

> It is not just bare theories and rules [*theoremata* and *kanones*]; it deals with things and has real beings as a kind of material for its activity; it approaches them methodically and possesses real

[29] *Ennead* V.8 [31].5f, citing Plato, *Symposium* 215b; see also *Ennead* IV.3 [27].11. It is commonly suggested that Plotinus is here mistaking the nature of hieroglyphs, but there is good reason, despite later oversimplifications or inventions (on which see Erik Hornung, *The Secret Lore of Egypt: its Impact on the West*, translated by David Lorton (Ithaca and London: Cornell University Press, 2001)), to think that he got it right. 'The Egyptians do not hesitate to call hieroglyphs "gods", and even to equate individual signs in the script with particular gods; it is quite in keeping with their views to see images of the gods as signs in a metalanguage. As is true of every Egyptian hieroglyph, they are more than just ciphers or lifeless symbols; the god can inhabit them, his cult image will normally be in the same form, and his priests may assume his role by wearing animal masks' (Erik Hornung, *Conceptions of God in Ancient Egypt: the One and the Many*, translated by John Baines (Ithaca: Cornell University Press, 1982), 124).

[30] *Ennead* III.7 [45].4. See also *Ennead* V.5 [32].1: 'They [*sc*. truth and the intelligible] are certainly not "premises" or "axioms" or "expressions"; for then they would only say something about other things and would not be the things themselves, as when [one says] "Justice is beautiful", though justice and beauty are different [from the words used].'

[31] *Ennead* 4.3 [27].30, 12–6.

> things along with its theories … So it does not know about propositions (*protasis*) – they are just letters – but in knowing the truth it knows what they call propositions, and in general it knows the movements of the soul, what it affirms and what it denies, and whether it affirms the same thing as it denies or something else, and if things are different from each other or the same; whatever is submitted to it it perceives by directing intuition, as sense perception also does, but it hands over petty precisions of speech to another discipline which finds satisfaction in them.[32]

So Plotinus' conception of knowledge strictly so-called is rather of a direct acquaintance with reality than anything that requires a language. And this is actually quite plausible. Most of us can recall a few scientific formulae, but we do not demonstrate our understanding merely by repeating the words, nor even by translating them into our mother tongue. We begin to understand them when we can *see* them, and the little flutter of delight that accompanies such dawning comprehension is a hint of the spirit that more dedicated lovers feel: 'this is the spirit that Beauty must ever induce, wonderment and a delicious trouble, longing and love and a trembling that is all delight.'[33] Can we even say that the *route* to knowledge must be dialectical, and so propositional? Intellect is accessible to us (as humans) because our souls have not descended too far from it (or rather, that is what it *means* that our souls have not descended too far). Our souls retain a *logos* of the intellect, and can reason their way back to it, even if they abandon discursive reason at the end.

> It [*sc*. dialectic] stops wandering about the world of sense and settles itself down in the world of intellect, and there it occupies itself, casting off falsehood and feeding the soul in what Plato calls 'the plain of truth', using his method of division to distinguish the Forms, and to determine the essential nature of each thing, and to find the primary kinds, and weaving together by the intellect all that issues from these primary kinds, till it has traversed the whole intelligible world; then it resolves again the structure of that world into its parts, and comes back to its starting point; and then, keeping quiet (for it is quiet insofar as it is present There) it busies itself no more, but contemplates (*blepei*), having arrived at unity. It leaves what is called logical

[32] *Ennead* I.3 [20].5.
[33] *Ennead* I.6 [1].4, McKenna's translation.

> activity, about propositions and syllogisms, to another art, as it might leave knowing how to write.[34]

Though dialectic is a favoured route to reason, it does not follow that higher souls are always thinking propositionally. Nor is it the only possible way: lovers and musicians have their own directions.[35] As I have already observed, Plotinus also advocates the use of wordless imagery to awaken intellect. Words come later.

> They [*sc*. the stars] always see [God]; and while they see him it is surely not possible for them to say that they have seen him: this would be something that would happen to those who have ceased to see.[36]

The God that the star-gods see, in this context, is the Intellect itself. The route that we must follow to that vision is itself one that we can talk about once we have traversed it – but the travel itself need not be – perhaps must not be – accompanied by any verbal monologue.

> The soul runs over all truths, and all the same shuns the truths we know if someone tries to express them in words and discursive thought; for discursive thought, in order to express anything in words, has to consider one thing after another: this is the method of description; but how can one describe the absolutely simple [the One]? But it is enough if the intellect comes into contact with it; but when it has done so, while the contact lasts, it is absolutely impossible, nor has it time, to speak; but it is afterwards that it is able to reason about it.[37]

In other words, we *can* talk about the truth, but *seeing* the truth is positively inimical to speech. 'In proportion as the confidence is clearer, the contemplation is quieter … [and] what it utters, it utters because of its deficiency, with a view to examining it, trying to learn thoroughly what it possesses'.[38] Nor is speech the only or the best route upwards to the truth. We do not even need speech to understand.

[34] *Ennead* I.3 [20].4, citing Plato, *Phaedrus* 248b6.
[35] See *Ennead* I.3 [20].1f.
[36] *Ennead* IV.4 [28].7; see Edward W. Warren, 'Consciousness in Plotinus', *Phronesis* **9** (1964), 83–97: even here-now we are only conscious that we were reading when we are no longer absorbed in the content of what we were reading.
[37] *Ennead* V.3 [49].17.
[38] *Ennead* III.8. [30].6, 14 & 27–9.

> Nor do I think that we should suppose that they use speech in the intelligible world, and altogether, even if they have bodies in heaven, there would be none of that talk there which they engage in here because of needs or over doubtful and disputed points; but as they do everything they do in order and according to nature they would not give orders or advice and would know by intuition (*sunesis*) what passes from one to another. For here below, too, we can know many things by the look in people's eyes when they are silent; but there all their body is clear and pure and each is like an eye, and nothing is hidden or feigned, but before one speaks to another that other has seen and understood.[39]

Speech may actually disguise our meaning – so 'let [one who wishes to understand] abandon the verbal signification and grasp the meaning of what is being said':[40] a plea that any author must occasionally murmur.

Compare and contrast Carruthers' remark:

> To say that creative thinking is done wordlessly may only mean that it is done non-consciously. Not for nothing have poets traditionally prayed to the Muses for inspiration; for we often have no idea where our genuinely novel ideas come from, nor is there much that we can do intentionally in order to get them. Sometimes a relaxing environment can help – a hot bath, a daydream, or a good night's sleep. But when ideas do come, they seem to us to come of their own accord, often with no discernible history.[41]

Carruthers' actual arguments for the 'non-consciousness' of wordless 'thought' are feeble. But Plotinus too suggests, as I remarked earlier, that we need not be aware of everything that affects us. His example (that of the great sea-beasts) is actually used to suggest that you and I *might* after all be a single soul even though neither of us, here and now, as the soul-body composites we currently are, shares the

[39] *Ennead* IV.3 [27].18; see also V.8 [31].4, and II.3 [52].7. The phrase is also used by J. K. Kadowaki, *Zen and the Bible*, translated by J. Rieck (London: Routledge & Kegan Paul, 1980), 33: working on Zen *koans* has the object of extending enlightenment from the mind's eye through the whole body, so that 'the whole body is an eye'. Kadowacki (*Ibid*., 13, 49) also emphasises the body language that precedes 'oral speech'.

[40] *Ennead* VI.4 [22].2; see also IV.8.9.

[41] Peter Carruthers, *Language, Thought and Consciousness: an Essay in Philosophical Psychology, op. cit*., 59; see also 138f.

other's perceptions. It is the nature of soul, after all, to be diverse: to experience things severally. Even our individual soul is not always fully transparent to itself – in fact it hardly ever is. We don't know who we are, nor what we know.[42] 'We are always intellectually active but do not always apprehend our activity.'[43] And it is not even clear that we always should, or that we should for always. *Self*-consciousness may be a foolish distraction from the consciousness of something better,[44] though we may need for a while to enlarge our awareness of what lies within. The goal of the Plotinian intellectual is to join 'the dance of immortal love', but it is a mark of the good dancer that she is not conscious of what she does.

> The dancer's intention looks elsewhere; his [*sic*] limbs are affected in accordance with the dance and serve the dance, and help to make it perfect and complete; and the connoisseur of ballet can say that to fit a particular figure one limb is raised, and other bent together, one is hidden, another degraded; the dancer does not choose to make these movements for no reason, but each part of him as he performs the dance has its necessary position in the dancing of the whole body.[45]

Our partial intellects *may* know themselves, as one such intellect amongst many, one of the multitudinous communities that are Plotinus' intelligible cosmos.[46] But the proper focus of intellect is the One, and the proper focus of soul is intellect itself. *Self*-consciousness in the ordinary sense is closer to the error made by

[42] *Ennead* VI.7 [38].41.

[43] *Ennead* IV.3 [27] .30,15. See Richard E. Aquila, 'On Plotinus and the Togetherness of Consciousness', *Journal of the History of Philosophy* **30** (1992), 7–32.

[44] *Ennead* I.4 [46].10: a reader is not conscious of the act of reading.

[45] *Ennead* IV.4 [28].33; see also VI.9 [9].38.

[46] In Hilary Armstrong's words: 'Plotinus' divine mind is not just a mind knowing a lot of eternal objects. It is an organic living community of interpenetrating beings which are at once Forms and intelligences, all "awake and alive", in which every part thinks and therefore is the whole; so that all are one mind and yet each retains its distinct individuality without which the whole would be impoverished. And this mind-world is the region where our own mind, illumined by the divine intellect finds its true self and lives its own life, its proper home and the penultimate stage on its journey, from which it is taken up to union with the Good' (A. H. Armstrong and R. A. Markus, *Christian Faith and Greek Philosophy* (London: Darton, Longman & Todd, 1960), 27).

Narcissus:[47] of focusing outside our own activity upon some product of that activity. Or else it lies in thinking how we might seem to others. Self-consciousness enfeebles activity.[48] 'Consciousness is more of a memory than a presence. It is inexorably tangled up in time. All it can give us is images, which it tries to fixate by expressing them in language.'[49] Better to call upon a Muse. Better, certainly, to call upon a Muse than to expect an answer from mere Matter.

Notoriously, Matter is almost as difficult to conceptualize as the One itself. None of us ever encounters Matter directly, but only the golden chains that bind it.[50] Soul (both the Soul of the All, and the Soul of each of us) creates the living world: without it there would only be 'the darkness of matter and non-existence, and "what the gods hate"'.[51] One view has been that it is

> By names that the world was called,
> Out of the empty air,
> With names was built and walled,
> Line and circle and square, dust and emerald;
> Snatched from deceiving death
> By the articulate breath.[52]

And without those names, 'animals' are – it has been thought – contained within an 'unchanging Here'. Plotinus' position is closer to William James':

> We may, if we like, by our reasonings, unwind things back to that black and jointless continuity of space and moving clouds of swarming atoms which science calls the only real world. But all the while the world we feel and live in will be that which our ancestors and we, by slowly cumulating strokes of choice, have extricated out of this, like sculptors, by simply rejecting certain portions of the given stuff. Other sculptors, other statues from

[47] See *Ennead* I.6 [1].8.

[48] *Ennead* I.4 [46].10.

[49] Pierre Hadot, *Plotinus, or The Simplicity of Vision*, translated by Michael Chase (Chicago: Chicago University Press, 1993), 32. Amusingly, the notion that consciousness is a report of what *has* happened also turns up in Daniel Dennett's *Consciousness Explained* as a modern insight.

[50] *Ennead* I.8.15; cf. *Ennead* I.6.9, on the golden screen or advance guard that is placed between us and the Good.

[51] *Ennead* V.1.2, after *Iliad* 20.65.

[52] Edwin Muir, 'The Animals' in *Collected Poems* (London: Faber, 1960), 208.

> the same stone! Other minds, other worlds from the same monotonous and inexpressive chaos! My world is but one in a million alike embedded, alike real to those who may abstract them. How different must be the worlds in the consciousness of ant, cuttlefish or crab![53]

The worlds we make – that all souls make – need not rely on names, nor any other words. Which is just as well: if we had to rely on words to make our worlds, we would never even have acquired the words to make them. It is not naming, but our *attention* that calls the worlds to life, even if it is by song – a wordless song – that this attention is attracted.

'Looking out of the window of your house or monad, you do not see words but things',[54] even though we who have learnt to speak find that our 'soul permanently rustles with words that are gradually taking over the density of things and their representation'. Calling a halt to that incessant chatter – and especially to that lumbering giant, the Ego – is a necessary step before we can even see what is straightforwardly in front of us.

3. How to Lose Your Mind[55]

But the goal is not only wordless, and self-forgetful. It is also in another sense quite mindless. We may reason our way towards a proper understanding and simultaneous recognition that what we then realize has been there all along, but vision transcends both words and sober thought.

> Intellect also, then, has one power for thinking, by which it looks at the things in itself, and one by which it looks at what transcends it by a direct awareness and reception, by which also before it saw only, and by seeing acquired intellect and is one.

[53] William James, *The Principles of Psychology* (New York: Macmillan, 1890), Vol.1, 288f.

[54] Cornelis Verhoeven, 'Wordless', translated by Jo Verhoeven in *Wordlessness*, edited by Bart Verschaffel and Mark Verminck (Antwerp: Lilliput Press, 1993), 36–42. My thanks to Panayiota Vassilopoulou for this reference.

[55] Some of the material in this section was presented at an ISNS conference at the University of Liverpool in the summer of 2004. See also 'Conclusion' in Panayiota Vassilopoulou and Stephen R. L. Clark (eds), *Late Antique Epistemology: Other Ways to Truth* (Basingstoke: Palgrave Macmillan, 2009), 289–301.

> And that first one is the contemplation of Intellect in its right mind, and the other is Intellect in love, when it goes out of its mind 'drunk with the nectar'; then it falls in love, simplified into happiness by having its fill, and it is better for it to be drunk with a drunkenness like this than to be more respectably sober.[56]

What sort of drunkenness is this? And how can it be better to be *drunk*? As so often, Plotinus deliberately subverts more usual metaphors, and may be signaling this in his choice of *nectar* rather than wine as the intoxicating agent.[57] For most philosophers, getting drunk is losing one's mind and morals. So Philo of Alexandria: allegorizing the story of Lot's daughters, he suggests that 'they made their father drink Wine' means that 'they brought complete insensibility on the mind, so that it fancied itself competent by its own abilities to judge what was expedient, and to assent to all sorts of apparent facts, as if they really had solid truth in them'.[58]

And then there are hangovers. Wine creates as a temporary condition something of the manic-depressive temperament, a neglect of ordinary concerns and a conviction that the world, somehow, is ours, quickly followed by despair. 'Intellect is our king [and] we too are kings when we are in accord with it',[59] but that may sound (at first) a little too much like the drunken conviction that 'Glasgow belongs to me'! According to Chrysippus virtue may be lost in consequence of drunkenness or melancholy,[60] and melancholia has, notoriously, been associated with academic genius since Aristotle.[61]

[56] *Ennead* VI.7 [38].35. This is closer to the thesis of *The Cloud of Unknowing* than to that of Ps-Dionysius. 'For Denys the divine darkness lies beyond the farthest effort of the mind, and it is the mind (the *nous*) that enters it: for the author of the *Cloud*, we enter the cloud of unknowing when we renounce the activity of the mind and rely solely on "the loving power" of the soul': Andrew Louth, *Denys the Areopagite* (London: Continuum, 1989), 125.

[57] *Ennead* III.5 [50].7: wine had not been invented at the birthday celebrations for Aphrodite, at which Poros got drunk and was seduced by Penia (after Plato's weird story in *The Symposium*: that is, this allegorical event 'precedes' material reality).

[58] Philo, *De Ebrietate* 41, citing *Genesis 19.33*.

[59] *Ennead* V.3.3, 46ff.

[60] Diogenes Laertius, *Lives of the Philosophers* 7.65.

[61] R. Klibansky, E. Panofsky, F. Saxl, *Saturn and Melancholy* (Edinburgh: Nelson, 1964).

> Wine in large quantity produces in men much the same characteristics which we attribute to the melancholic, and as it is being drunk it fashions various characters, for instance irritable, benevolent, compassionate or reckless ones.... We are often in a state of grieving, but could not say why, while at other times we feel cheerful without apparent reason.[62]

It is in this context also that Philo warns against rationalist conceit: 'If we mistakenly trust our private reasonings we shall construct and build the city of the mind that destroys the truth... For this reason one who has had a dream finds on awakening that all the movement and exertions of the foolish men are dreams devoid of reality. Indeed mind itself was found to be a dream'.[63] He goes on to mock philosophical pretensions, and philosophers' undue willingness to trust the fallible instruments of reason and experience, pointing out that 'about these very things, and about the different ways of life, and about the ends to which all actions ought to be referred, and about ten thousand other things which logical, and moral, and natural philosophy comprehends, there have been an unspeakable number of discussions, as to which, up to the present time, there is no agreement whatever among all these philosophers who have examined into such subjects'.[64] Philo was as insistent as any modern that our imaginings and reasonings need to be checked, that we need, somehow, to 'get in touch' with the real.[65] And just as a reminder that these matters are not without their practical and political dimension:

> The moral collapse of Vietnam was scarcely caused by an overdose of objective consciousness about what we were doing. It consisted of the failure to expand consciousness beyond mere instrumental tasks to the practical and banal significance of our

[62] Aristotle, *Problemata* 30. He adds that wine is also aphrodisiac, and can sometimes improve a poet (unless the remark is ironical): 'Maracus the Syracusan was a better poet when he was out of his mind'.

[63] Philo, *Legum Allegoriarum* 3.228f. in *Philo of Alexandria: the Contemplative Life, The Giants, and Selection*s, translated by D. Winston (London: SPCK, 1981), 151.

[64] Philo, *De Ebrietate*, 48.

[65] See also the Hermetic Corpus 1.27 (*Hermetica*, translated by Brian P. Copenhaver (Cambridge: Cambridge University Press, 1992), 6): 'People, earthborn men, you who have surrendered yourself to drunkenness and sleep and ignorance of god, make yourselves sober and end your drunken sickness, for you are bewitched in unreasoning sleep' (see also 7.1, *Ibid.*, 24).

> national goals and policies. We kept the war going in Vietnam because our consciousness was mystified by symbols of patriotism, dreams of glory, unyielding pride, and visions of empire. In mood we were exactly what the counter-culture people want us to become. We imagined we were menaced by slant-eyed devils and worthless little yellow men; we enthralled ourselves with visions of our own ineffable majesty. In short, we were stoned.[66]

'Mystics', of course, may tell a different story about drink (including Philo, on other occasions). According to the Chaldaean Oracles, for example, the soul once freed from the body and filled with noetic light 'glories in the harmony with which it is drunken'.[67] Philo also speaks approvingly of the '*sober* intoxication' that possesses the mind 'when it has transcended all sensible substance… on beholding in [the intelligible] realm beauties beyond measure, the patterns and originals of the sensible things in the world below'.[68]

> If then, my soul, a yearning comes upon you to inherit the divine goods, abandon not only your land, that is, the body; your kinsfolk, that is, the senses; your father's house, that is, speech, but escape also your own self and stand aside from yourself, like persons possessed and corybants seized by Bacchic frenzy and carried away by some kind of prophetic inspiration.[69]

That wine, and yet more potent drugs, have this ambiguous reputation is hardly surprising. Back in the days of the 'counterculture' LSD might possibly produce euphoric insights, or hellish hallucinations. Respectable people may plausibly consider all such drugs are dangerous, while also, privately, acknowledging that those strange insights – and even the hellish hallucinations – may sometimes be fruitfully life-changing.[70] Mainstream psychiatry suspects all

[66] Marvin Harris, *Cows, Wars, Pigs and Witches* (New York: Vintage Books, 1989; 1st published 1974), 266.

[67] Hans Lewy, *Chaldaean Oracles and Theurgy* (Le Caire: L'Institut Français d'archéologic orientale, 1956), 420. See also Psellus Comm., 1137A (Kroll 48): Lewy, 198n: 'having nothing mortal she is utterly drunk'.

[68] Philo, *De Opificio* Mundi, 70 in D. Winston (tr.), *Philo of Alexandria: the Contemplative Life, The Giants, and Selections, op. cit.*, 173. See further Hans Lewy, *Sobria Ebrietas* (Giessen: A. Töppelmann, 1929).

[69] Philo, *Quis Rerum Divinarum Heres Sit*, 68f in D. Winston (tr.), *op. cit.*, 169.

[70] See John E. Nelson, *Healing the Split* (New York: SUNY Press, 1994), 147–52 for a balanced and well-informed discussion of the issues.

claims to 'higher' consciousness, ignoring what may otherwise seem evident distinctions between psychosis and transcendence (as did the anti-psychiatric movement of the 60s and 70s, with opposite evaluations).[71] I prefer to suppose that there are insights to be had from a long tradition of consciousness-altering practices and theories. Simply calling those insights 'mystical' is unhelpful. It may be that Plotinus is talking about just this sort of ecstasy – but what is 'ecstasy'? What is it even to be 'drunk', whether with love, liquor or sudden madness? 'Drunkenness', according to the anthropologist, 'also expresses culture in so far as it always takes one form of a highly patterned, learned comportment which varies from one culture to another: pink elephants in one region, green snakes in another.... Drinking is essentially a social act, performed in a recognized social context'.[72] Even within a single culture, there are gloomy drunks, and sleepy ones, and violent ones; drunks who dance and sing, and drunks who tell secrets to passing strangers.

The likeliest association for Plotinus is with the passion of love – by which I mean not *charity* but, bluntly, sex.[73] Sexual experience must simultaneously be a real encounter – no-one, Plotinus says, would prefer merely to *imagine* the presence of the beloved[74] – and one that lets us put aside our selves.[75] The beauty that such drunken lovers see has 'penetrated through the whole of their soul', and they are not simply spectators – 'as if someone possessed by a god, taken over by Phoebus or one of the Muses, could bring about the vision of the god in himself, if he had the power to look at the god in himself'.[76] And even this is to be surpassed:

[71] *Ibid*., 358–9.

[72] Mary Douglas, 'A distinctive anthropological perspective' in Mary Douglas (ed.), *Constructive Drinking: Perspectives on Drink from Anthropology* (Cambridge: Cambridge University Press, 1987), 3–15:4.

[73] *Ennead* VI.7[38].34, 14. See John Rist, *Eros and Psyche: Studies in Plato, Plotinus and Origen* (Toronto: University of Toronto Press, 1964), 77–9: 99. See also Zeke Mazur, 'Having Sex with the One: Erotic Mysticism in Plotinus and the Problem of Metaphor' in P. Vassilopoulou and S. R. L. Clark (eds), *Late Antique Epistemology*, *op. cit*., 67–83.

[74] *Ennead* VI.7 [38].26, 21f.

[75] I should add – merely to avoid misunderstanding - that Plotinus reckons that 'if [lovers] remain chaste there is no error in their intimacy with the beauty here below, but it is error to fall away into sexual intercourse': *Ennead* III.5 [50].1, 36f.

[76] *Ennead* V.8 [31].10, after Plato, *Phaedrus* 246eff.

> As if carried away or possessed by a god, in a quiet solitude and a state of calm, not turning away anywhere in his being and not busy about himself, altogether at rest and having become a kind of rest. He had no thought of beauties but had already run up beyond beauty and gone beyond the choir of virtues, like a man who enters into the sanctuary and leaves behind the statues in the outer shrine. …. They are secondary objects of contemplation. But that other, perhaps, was not a contemplation but another kind of seeing, a being out of oneself (*ekstasis*) and simplifying and giving oneself over and pressing towards contact and rest and a sustained thought leading to adaptation (*perinoesis pros epharmogen*), if one is going to contemplate what is in the sanctuary.[77]

Remember that it is those who have *ceased* to see, who can say what they have seen.[78] Our goal is self-forgetfulness, 'on the far side of thought and being'.

4. There and Back Again

Materialists, even the sanely inconsistent ones, are bound to suppose that the first reality of things, the world which was before us and will, presumably, outlast us all, is wholly unreflective. There are no privileged places, times or scales in that 'real world'. In the ages – neither short nor long – before there were living creatures there was nothing either good or bad, large or small, long-lasting or short-lived. In the ages before there were time-binding creatures there were no narrative histories or temporal identities: everything arose from everything by whatever 'laws' somehow subsisted 'there'. 'There', on the far side of any human or any animal experience, everything is altogether, and it is only our strange error to suppose that things are any different 'now'. Somehow – and I shall make no effort here to consider how this has come about – we are equipped to notice and acknowledge and sometimes even *prefer* that vision of things to the merely egoistical, parochial and fleeting consciousness we actually inhabit. Reason, on Plotinus' account, is not to be valued because we can, with its help, perform socially useful actions, nor yet because we can, with its help, win more food, water or shelter than less 'clever' creatures.[79] It is to

[77] *Ennead* VI.9 [9].11.
[78] *Ennead* IV.4 [28].7.
[79] Actually, as he points out, non-rational creatures manage very well: *Ennead* I.4 [46].2.31–43.

be valued – and insight is especially to be valued – because it is the presence in us of beauty. 'Intellect is our king. But we too are kings when we are in accord with it; we can be in accord with it in two ways, either by having something like its writing written in us like laws, or by being as if filled with it and able to see it and be aware of it as present'.[80] Either we simply recite the creed to ourselves, or we somehow, momentarily, live it. But of course in *living* it we are still untrue to it: even the clearest and least parochial glimpse of truth is still untrue, because there is really no such split between observed and the observer. We falsify the world by looking at it.

> The beauty of a landscape just at the moment when no-one is looking at it, absolutely nobody … To see a landscape such as it is when I am not there. When I am anywhere, I pollute the silence of earth and sky with my breathing and the beating of my heart.[81]

Only in *losing* consciousness can we be true – or rather, since *truth* is, exactly, what is *not* forgotten,[82] in losing consciousness we transcend the truth. We're only in the right when we've forgotten how. *Intellect* is king, but not *my* intellect. Nature, the barest trace of soul in the material, temporal world, works in silence, in contemplation: 'what comes into being is what [she] sees in [her] silence'.[83]

[80] *Ennead* V.3 [49].3, 46ff.

[81] Simone Weil, *Notebooks*, translated by A. Wills (London: Routledge & Kegan Paul, 1956), Vol.2, 423. This thought is not to be taken lightly: it is difficult not to see in it one aspect of Weil's fatal anorexia, but her insights should not be ignored merely because, as it so often does, 'the disease' (that is, the demon) took advantage of them. When she wrote that 'I cannot conceive of the necessity for God to love me, when I feel so clearly that even with human beings affection for me can only be a mistake. But I can easily imagine that he loves that perspective of creation which can only be seen from the point where I am… I must withdraw so that he can see it. I must withdraw so that God may make contact with the beings whom chance places in my path and whom he loves' (Simone Weil, *Gravity and Grace* (Lincoln, NE: University of Nebraska Press, 1952), 88), the disease was speaking. How could it be that God loved everything but Weil? See also John Lippitt, 'True self-love and true self-sacrifice', *International Journal of Philosophy of Religion* **64–5** (2009).

[82] The punning relationship of *aletheia*, truth, and *lethe*, forgetfulness, is founded on a false etymology, but still has some force: see my *God's World and the Great Awakening* (Oxford: Clarendon Press, 1991), 48–54.

[83] *Ennead* III.8 [30].4. 5–6.

The transcendent cause of everything, including Nature, is not itself a thing.

> It is therefore truly ineffable: for whatever you say about it you will always be speaking of a 'something'. But 'beyond all things and beyond the supreme majesty of Intellect' is the only one of all the ways of speaking of it which is true; it is not its name, but says that it is not one of all things and 'has no name', because we can say nothing of it: we can only try, as far as possible, to make signs to ourselves about it.[84]

What we cannot talk about we cannot reason about either. Notoriously, this makes it difficult to understand what Plotinus meant.

> The perplexity arises especially because our awareness of the One is not by way of reasoned knowledge (*episteme*) or of intellectual perception (*noesis*) as with other intelligible things, but by way of a presence superior to knowledge. The soul experiences its falling away from being one and is not altogether one when it has reasoned knowledge of anything; for reasoned knowledge is a rational process (*logos*), and a rational process is many. The soul therefore goes past the One and falls into number and multiplicity. One must therefore run up above knowledge and in no way depart from being one, but one must depart from knowledge and things known, and from every other, even beautiful, object of vision.… Therefore, Plato says, 'it cannot be spoken or written', but we speak and write impelling towards it and wakening from reasonings to the vision of it, as if showing the way to one who wants to have a view of something.[85]

The One itself can neither speak nor know Itself.

> If [the intellect] directed its gaze to a single object without parts, it would be without thought or word (*elogethe*): for what would it have to say about or to understand? For if the absolutely partless had to speak itself, it must, first of all, say what it is not; so that in this way too it would be many in order to be one. Then when it says 'I am this', if it means something other than itself by 'this', it

[84] *Ennead* V.3 [49].13. That the One is not a thing is a doctrine ably discussed by Eric D. Perl in *Theophany: the Neoplatonic Philosophy of Dionysius the Areopagite* (New York: SUNY Press, 2007). On the apophatic tradition, see Michael Anthony Sells, *Mystical Languages of Unsaying* (Chicago: University of Chicago Press, 1994).

[85] *Ennead* VI.9 [9].4, quoting Plato, *Letter* VII.241c5.

> will be telling a lie; but if it is speaking of some incidental property of itself, it will be saying that it is many or saying 'am am' or 'I I'.[86]

The One is the transcendent cause of all that is, seen and unseen – and the chief reason why Plotinus is so often regarded as a 'mystic'. Actually, there is another and more naturalistic interpretation: Plotinus is speaking of the various modes of consciousness available to us, and suggesting that there is a radically *simple* mode of a kind familiar in Buddhist practice. Soul experiences things in a linear and fragmentary way. Intellect experiences them as an ordered, beautiful whole. But even beyond Intellect lies a mode of consciousness that is not consciousness *of* anything at all: the bare light of awareness, without any doubleness of subject and object. Is that bare awareness really *awareness* at all? If intellect, object and thinking were the same, all would disappear[87] – and do, into the One. Or contrariwise, it is from the Nameless that Heaven and Earth spring.[88]

So are materialists and Plotinians, in the end, united? So David Hume, through his character Cleanthes, appositely enquires 'how … mystics, who maintain the absolute incomprehensibility of the Deity, differ from Sceptics or Atheists, who assert, that the first cause of all is unknown and unintelligible?'[89] Interestingly, there are 'materialists' who seek to solve 'the Hard Problem' by insisting that since we don't know what 'matter' is we can't rule out the possibility that it has properties sufficient to make 'awareness' an intelligible outcome. 'Matter' is something we know not what! How then indeed does materialism differ from mysticism? How do the star-gods differ from Plotinian plants, who have gone so far into the material as to forget themselves entirely?[90] How does the One, hidden behind its golden veil, differ from Matter, bound in golden chains?[91] We cannot speak accurately of either, and both, in their way, are the *potential*, the starting point, of everything. It is intriguing that, on Hornung's interpretation, 'for the Egyptians the world emerges from the one, because the non-existent is one', and

[86] *Ennead* V.3 [49].10; see also VI.7 [38].35.

[87] *Ennead* VI.7 [38].41; see also *Ennead* V.1 [10].4, 38: 'if you take away otherness, it will become one and remain silent'.

[88] *Tao Te Ching*, ch.1.

[89] David Hume, *Dialogues Concerning Natural Religion (*1779*)*, pt.4: *Hume on Religion*, edited by R. Wollheim (London: Fontana, 1963), 131.

[90] *Ennead* V.2 [11].2: plants embody the most audacious and stupid part or sort of soul!

[91] Compare *Enneads* I.6 [1].9, 39 and I.8 [51].15, 25.

that non-existent is also the nothingness, the void, that surrounds the cosmos.[92] But that is another story.[93]

Phenomenologically at least, Plotinus, as it seems to me, has the superior insight: there is a difference, at least for us, in the direction that we face, whether towards the stuff of mere sensation,[94] or towards the One, by which we encounter beauty. There is a difference between two sorts of drunkenness: are we to dull our minds with wine, so to speak, or drink the heavenly nectar? Do we prefer the vegetative dream, or the stellar? The answer is not to be deduced from any prior reasoning, though we may be led, by reason, to a point where it can be 'seen'.

> It is there that one lets all study go; up to a point one has been led along and settled firmly in beauty and as far as this one thinks that in which one is, but is carried out of it by a kind of swell and sees suddenly, not seeing how, but the vision fills his eyes with light and does not make him see something else by it, but the light itself is what he sees.[95]

Seeing and being that light is no longer to be conscious of a self distinct from it. The memory of that light is what gives us such selfhood as we possess.

University of Liverpool

[92] See also Ps-Dionysius, *The Complete Works*, translated by Colm Luibheid (London: SPCK, 1987), 73: *The Divine Names* 697A: 'Its nature, unconfined by form, is the creator of all form. In it is nonbeing really an excess of being. It is not *a* life, but is, rather, superabundant Life. It is not *a* mind, but is superabundant Wisdom.... And one might even say that nonbeing itself longs for the Good which is above all being.'

[93] E. Hornung, *Conceptions of God in Ancient Egypt, op. cit.*, 253.

[94] This collapse of boundaries, 'the death of the "I" as knower, chooser, and doer', is the sinister image of the Plotinian ascent: see John E. Nelson, *Healing the Split*, *op. cit.*, 43. In a way the psychotic is correct: reality is larger than we know, and full of demons. But if Plotinus is correct, so also was Housman, in his more optimistic mode: 'The world is round, so travellers tell,/ And straight though reach the track,/ Trudge on, trudge on, 'twill all be well,/ The way will guide one back' (A. E. Housman, 'A Shropshire Lad' 36 in *Collected Poems* (Harmondsworth: Penguin, 1956; 1st published 1939), 65).

[95] *Ennead* VI.7 [38].36.

The Road to Substance Dualism

GEOFFREY MADELL

Abstract
The common materialist view that a functional account of intentionality will eventually be produced is rejected, as is the notion that intentional states are multiply realisable. It is argued also that, contrary to what many materialists have held, the causation of behaviour by intentional states rules out the possibility of a complete explanation of human behaviour in physical terms, and that this points to substance dualism. Kant's criticism of the Cartesian self as a substance, endorsed by P. F. Strawson, rests on a misinterpretation of Descartes. The so-called 'causal pairing problem', which Kim sees to be the crucial objection to substance dualism, is examined, and Kim's arguments are rejected.

1. Why a Functional Account of Intentionality is not Possible

> I don't see principled objections to a functional account of intentionality. Let me say here that it seems to me inconceivable that a possible world exists that is an exact physical duplicate of this world but lacking wholly in intentionality.[1]

This claim of Kim's expresses a view which is widely held. The standard materialist view seems to be that the real problem for the physicalist (the 'hard' problem) is presented by 'qualia' rather than by intentionality itself. For reasons I shall explain, I find this position quite incredible. But first, let us note that Kim's supporting consideration expressed in the second sentence in fact offers no support for a physicalist conception of intentionality at all. It is surely obvious that an interactionist dualist will accept that there cannot be a duplicate world lacking intentionality, since to remove the intentional causes of behaviour will result in a world which cannot be a duplicate of this one. Kim offers another supporting consideration for his view, which runs as follows:

> Consider a population of creatures … that are functionally and behaviourally indistinguishable from us … If all this is the

[1] J. Kim, *Mind in a Physical World* (Cambridge, Mass.: The MIT Press, 1998), 101.

doi:10.1017/S1358246110000111

> case, it would be incoherent to withhold states like belief, desire, knowledge, action and intention from these creatures.[2]

This equally offers no support for a physicalist position, since, once again, it is something with which any interactionist dualist can readily agree. If these creatures really are behaviourally indistinguishable from us, we would ascribe to them much the same sort of inner life as we have ourselves, and which we ascribe to other people on the basis of their behaviour. It is difficult to understand how Kim could have supposed that either of these points offers support for physicalism.

Now let us look at Kim's claim that there are no principled objections to a functional construal of intentionality. The obvious objection to Kim's claim is that, while it is indeed possible to give what might look like a functional construal of intentional states in terms of their typical causal inputs and outputs, those inputs and outputs themselves are irreducibly mental. Indignation, to take one example, might be defined as that state which is brought about by the perception of a wrong or an insult, and leads in turn to the desire to protest. And gratitude might be defined as that state which is brought about by the recognition that one is the recipient of something which one values and which causes in its turn the desire to express one's thanks. And so on. But, of course, these states are irreducibly mental, or so I would claim.

Kim, by contrast, while allowing that no one has yet produced full functional definitions of believing, desiring and intending, and that it is 'perhaps unlikely that we will have such definitions any time soon',[3] sees no problem in the idea that such definitions will eventually be produced. We surely need to ask what such definitions could possibly look like, even in their most abstract form. One obvious obstacle that stands in the way of the possibility of such definitions is that an intentional state such as that of the experience of gratitude or indignation must involve the here-and-now or categorical direction of consciousness to its object: my thought is here and now directed to the benefit I have received and to a way of conveying to the benefactor how much I appreciate this benefit, or to the wrong I have suffered and to the need to make this clear. Such states are necessarily events in consciousness and necessarily involve the categorical, non-dispositional, directedness of consciousness to an object of thought. It is

[2] J. Kim, *Physicalism, or Something Near Enough* (Princeton, N.J.: Princeton University Press, 2005), 165.

[3] *Ibid*.

utterly unclear how this essential property of intentional states could be realised in an assembly of physical elements; at any rate, I know of no explanation of how this is possible And it is impossible to tell from Kim's discussion whether, in claiming that full functional definitions of states such as believing and desiring will eventually be found, he supposes that such definitions will show that these intentional states do not, after all, involve occurrent states of consciousness which are here-and-now directed to their objects, or, on the other hand, that this essential property of non-dispositional directedness to an object can be accommodated within the complete physicalist-functionalist story. The former option is, in my view, clearly false, and the second a bewildering claim.

2. Why Intentional Patterns are not in the Physical World

Dennett has argued for a position with which Kim may be in agreement, and examining it will bring the issue more into focus. In a departure from his earlier instrumentalist position, Dennett claims that there is a sense in which we can accept that intentional concepts pick out real patterns in the world. These patterns, however, are not visual patterns, but, one might say, *intellectual* patterns. To adapt one of his examples slightly, a range of chess games all played to a conclusion will reveal to the novice a completely haphazard collection of physical events. Eventually, however, one might expect that a certain pattern will begin to become evident: something about the way a selection of pieces of one colour is grouped around a particular piece of the other colour – checkmate. Similarly, the Martian peering through a telescope at the Superbowl game might eventually begin to understand what is going on, particularly if he adopts 'the intentional stance'. And in an earlier paper Dennett considers the Martian (again) looking at a stockbroker placing an order for 500 shares in General Motors, and goes on:

> But if the Martians do not see that indefinitely many *different* patterns of finger motions and vocal chord vibrations – even the motions of indefinitely many different individuals – could have been substituted for the actual particulars without perturbing the subsequent operation of the market, then they have failed to see a real pattern in the world they are observing.[4]

[4] D. Dennett, 'True Believers' in *The Intentional Stance* (Cambridge, Mass.: The MIT Press, 1987), 26.

Adopting the intentional stance in these cases, Dennett tells us, hugely increases one's predictive power, and enables one to discern the relevant patterns far more easily. These patterns, at any rate, are real, though perceiving them is not a matter of our being confronted with some easily perceptible visual pattern, but of discerning a high-grade 'intellectual' pattern, a pattern which may be difficult to detect through the surrounding 'noise'. And I suppose that this must also be Kim's view, though, for him, the noise is such as to have prevented us up to now from perceiving the underlying intellectual patterns, or algorithms. To discern this in any particular case is to be able to pick out the relevant paths through the physical world, the paths which constitute the different possible realisations of the intentional state.

The suggestion that 'the reality of intentional patterns', to use Dennett's phrase, can be accounted for in this way is deeply flawed. What primarily undermines the suggestion is that all Dennett's examples involve either games (chess, the Superbowl game) or an activity defined by certain rules or conventions (placing an order for shares in General Motors). In these examples, it is true that understanding what is going on is a matter of grasping the underlying pattern or algorithm, a matter of coming to understand the rules which govern the activity in question. And it is true that, while adopting 'the intentional stance' will help one to grasp what is going on much more easily (or so we can allow), it is not essential that one adopts this stance. The relevant paths through the physical world might be grasped without assuming intentions on the part of any imagined person.

But grasping that what one is witnessing is an expression of indignation or gratitude cannot be like this. There is no pattern to be discerned, not even the most high-grade or 'intellectual', no underlying rule or algorithm, something which might initially be discerned without presupposing any underlying intentional attitude. The only thing common to a range of possible expressions of indignation or gratitude or remorse is that they are all seen by us to be such expressions. Such understanding is not a matter of discerning some pattern which obtains in the physical world, a pattern which we might eventually perceive through the surrounding 'noise'. It is not an understanding which might be achieved without adopting the intentional stance at all. It is essentially intentional understanding, subjective or first-personal. It is understanding which is achieved initially by bringing the template of one's own conscious experience of what it is to have intentions and emotions to bear. And, to repeat a point made earlier, it is knowledge of intentional states as involving

the immediate, non-dispositional directedness of consciousness to its various objects. I take all this to show, contrary to what Kim and others suppose, that the expectation that we will eventually have a functionalist account of intentionality which is compatible with physicalism is totally misconceived.

3. Why Mental Causation Rules out Physicalism

Since intentional concepts do not pick out a pattern of pathways through the physical world, a pattern determined by some underlying algorithm, the question must arise, what is the relation between the intentional state and the physical events which are its expression? On the face of it, it is causal: my gratitude for the receipt of a gift, or my remorse for a wrong I have committed, leads me to act in a certain way. The course of events is determined by the nature or the logic of the intentional state. It seems to me incredible that a complete explanation in intentional terms of a certain course of behaviour might be paralleled by an equally complete explanation of the same course of behaviour couched in the terminology of the physical sciences, and one which involved no recourse to intentional notions. That would be simply a miraculous coincidence.

I take the term 'miraculous coincidence' from Adrian Cussins, who has forcefully described the challenge that the physicalist has to answer. However, his attempt to answer this challenge seems to me very puzzling. Here is a relevant passage:

> Would a mother hold her child close to the edge of the canyon so the child could see the view? She could count on her intention to hold the child tight, but neither folk psychology nor neurophysiology provides any assurance whatever that her neurophysiology will march in step with her intention. Isn't it a miracle that the predictions march in step? Of course not. *It is the nature of human cognition that this is how things are*. It is because humans have the cognitive nature that they have that their physiology meshes with folk psychology (Cussins' italics).[5]

I cannot see that this point does anything to meet the challenge to the physicalist that I outlined above. What is clear, surely, is that the mother's behaviour is determined by her love for her child and her

[5] A Cussins, 'The Limits of Pluralism' in David Charles and Kathleen Lennon (eds), *Reduction, Explanation and Realism* (Oxford: Clarendon Press, 1992), 198.

concern for its safety. These are the intentional states which determine her behaviour, and which thus determine the neurophysiological processes which underlie that behaviour. It is not, as Cussins claims, a matter of our physiology meshing with our psychology or our intentional states, but of our intentional states determining the course of our neurophysiological processes.

Imagine a mother who entirely lacks concern for her child, and is quite prepared to use her child as a way of making money. Standing at the edge of the canyon, she remembers that her holiday insurance covers accidental death, and her child then has an unfortunate accident. This mother has the same neurophysiology as the rest of us, but the neurophysiological processes that take place in this case are very different from those that take place in the loving mother, because the intentional states which determine her behaviour are very different.

I am equally puzzled by Galen Strawson's attempt to meet this problem. Here is the relevant passage:

> A decent stopping point in the mind body problem … would be to contemplate a fabulously detailed and exhaustive specification in neurological or particle-physics terms of the causation involved in a line of thought or a practical decision, and to feel no force in the objection that the availability of this specification showed that the mental was epiphenomenal or causally inefficacious …
>
> I think I have made it … One of the keys, I am sure, is to see that there is a fundamental component to the business of consciously entertaining and comprehending propositions that is just a matter of 'qualitative-experiential character' in every sense in which an experience of red is just a matter of qualitative-experiential character … It takes time, though.[6]

I cannot see how this passage does anything to meet the problem which Strawson outlines. To claim, as he does, that part of the answer is to recognise the 'qualitative-experiential character' of consciously entertaining a proposition is to emphasise something which the epiphenomenalist readily accepts. For to insist that consciously entertaining a thought has a qualitative character no more shows that it is not an epiphenomenon than pointing out that pain is a sensation, something having a qualitative character in consciousness, shows that the sensation is not a mere epiphenomenon. Epiphenomenalism readily accepts that such mental events do have

[6] G. Strawson, 'Panpsychism? Reply to Commentators', *Journal of Consciousness Studies* **13** (2006), 275–6.

a qualitative character, but that does not in any way indicate that they are not mere epiphenomena. Strawson's point does nothing to meet the central difficulty: if there is a complete explanation of human behaviour in terms of the categories available to the physical sciences, then the mental is epiphenomenal.

I have looked only at two responses to the problem of mental causation, and I will confess that, on looking back at much of what has been written, and continues to be written, on this issue, after a while I felt the life-force draining out of me and had to give up. So let me just say that I found nothing which threatened to overturn my view that the supposition that a complete explanation in intentional terms of one's acting from gratitude, jealousy, remorse, *etc*., can be paralleled by an equally complete explanation couched in the terms of the physical sciences would be to posit something utterly miraculous.

4. Why the Claim that Intentional States are Multiply Realisable Must be Rejected

One further suggestion that needs to be looked is the claim that what I have said would be utterly miraculous, *viz*. a complete explanation of a stretch of behaviour in intentional terms paralleled by an equally complete explanation of that same stretch of behaviour in physical terms, is not in the least miraculous, since intentional states and operations are realised in the purely physical world. We see this, for example, in the operations of a computer playing chess. Intentional concepts, we can allow, are irreducible to purely physical concepts, but that is only in the sense that the same intentional operation can be realised in different physical set-ups. An intentional activity such as that of playing chess might be undertaken by computers of different designs, for example.

This suggestion cannot stand. The game of chess is governed by certain rules, and, as I said above, grasping the underlying algorithm allows one to see the essential similarity common to a number of games played to a finish. And that is to say that that algorithm, those rules, can be computed. The algorithm determines what possible paths through the physical world are permissible. But the irreducibility of most intentional states is quite different from this. What counts as a possible expression of remorse or gratitude, of jealousy or indignation, is not something determined by an algorithm. To repeat: the only paths through the physical world which count as possible expressions of any of these intentional states are those which, in the first instance, our own experience of emotion enables us to see as such.

Papineau, by contrast, has indeed suggested that intentional states are multiply realisable in much the same way as, for example, the operation of thermostatic heaters is multiply realisable. There are thermostatic heaters of different designs, and that means that there is no single pathway to the end of heating water to a required temperature.[7] The suggested parallel is quite untenable, however. First, the various possible expressions of gratitude or indignation are not different pathways to some physically identifiable end, or an end defined by some set of rules or an algorithm, but to an end which our own subjective, first personal experience allows us to see as an end. Second, talk of intentional states being *realised* in some pathway through the physical world must be rejected, as I have already claimed. Various pathways through the physical world may indeed be *expressions* of an intentional state, but their being so is a matter of their being ways of achieving the end to which the thought of the subject is directed in that intrinsic, non-dispositional way which is an essential feature of intentionality, but which cannot be a feature of any aspect of the physical world. Simply to be confronted with a complex pattern of pathways through the physical world, *just* that, would leave one without any insight at all into the nature of an intentional state such as what it is to express sympathy for someone or what an expression of gratitude might be.

Recognising the fact that intentional states are causes of behaviour, and that the intentional cannot be reduced to the physical, or seen to be realised in sequences of physical events, leads inexorably to the conclusion that the principle of causal closure has to be abandoned. I think this in turn leads to interactionist substance dualism. I shall defend this claim by looking first at a well-known attempt to undermine the very notion of a mental substance, and then at an argument recently re-invoked by Kim which it is claimed shows that the notion of causal interaction between the physical and the non-physical is incoherent.

5. Why Kant's Rejection of the Notion of a Mental Substance is Misplaced

Here is a very well-known footnote from *The Critique of Pure Reason*:

> An elastic ball which impinges on another similar ball in a straight line communicates to the latter its whole motion … If,

[7] D. Papineau, 'Irreducibility and Teleology' in D. Charles and K. Lennon (eds), *Reduction, Explanation and Realism* (Oxford: Clarendon Press, 1992), 60.

> then in analogy with such bodies, we postulate substances such that the one communicates to the other representations together with the consciousness of them, we can conceive a whole series of substances of which the first transmits its state together with its consciousness to the second, the second its own state with that of the preceding substance to the third, and this in turn the states of all the preceding substances together with its own consciousness and with their consciousness to another. The last substance would then be conscious of all the states of the previously changed substances, as being its own states, because they would have been transferred to it together with the consciousness of them. And yet it would not have been one and the same person in all these states.[8]

A considerable number of commentators have endorsed this conclusion, though none more enthusiastically than P. F. Strawson, who writes:

> This line of attack could be pressed further than Kant presses it. Thus when the man … speaks, we could suggest that there are, perhaps, a thousand souls simultaneously thinking the thoughts his words express, having qualitatively indistinguishable experiences such as he, the man, would currently claim. How could the man persuade us that there was only one soul associated with his body? (How could the – or each – soul persuade itself of its uniqueness?)[9]

Strawson takes this to be 'the coup de grâce to Cartesianism'.[10] It is no such thing. It rests on a major misinterpretation of Descartes which quite undermines it. To see this, we need first to remind ourselves of Descartes' definition of 'substance': 'By *substance* we can understand nothing other than a thing which exists in such a way as to depend on no other thing for its existence'.[11] And it is absolutely clear that for Descartes thought itself is substantival. The idea of thought, that is, is an idea of that which can be understood completely, or as a complete thing. Descartes does indeed acknowledge

[8] I. Kant, *The Critique of Pure Reason*, translated by N. Kemp Smith (London: Macmillan, 1956), A 363–4, footnote.

[9] P. F. Strawson, *The Bounds of Sense* (London: Methuen, 1966), 168.

[10] *Ibid*.

[11] R. Descartes, *Principles of Philosophy*, I.51 in *Descartes: Selected Philosophical Writings*, translated by John Cottingham, Robert Stoothoff and Dugald Murdoch (Cambridge: Cambridge University Press, 1988), 177.

that there is a distinction between modes of thought on the one hand and the objects which have them, but insists that this distinction is merely a conceptual one. Here is the crucial passage:

> Finally, a *conceptual distinction* is a distinction between a substance and some attribute of that substance without which the substance is unintelligible … And in the case of all the modes of thought which we consider as being in objects, there is merely a conceptual distinction between the modes and the object which they are thought of as applying to … Thought and extension can be regarded as constituting the natures of intelligent substance and corporeal substance; they must then be considered as nothing else but thinking substance itself and extended substance itself – that is, as mind and body.[12]

You would think that nothing could be clearer than that: the distinction between all the various modes of thought on the one hand and the mind which has them on the other is merely a conceptual one. Thought, then, is not something that inheres in some underlying substratum. Indeed, if that were the case then we could not know of that substratum that it is not also the bearer of corporeal properties, and Descartes would have no argument at all for the real distinction between mind and body, for which he argues in the sixth *Meditation*.

What Kant appears to have done is to foist on Descartes an ontological distinction between thought on the one hand and the object in which thought inheres, ignoring Descartes' clear assertion that this distinction is merely a conceptual one. Kant's claim that the Cartesian view of mind would allow the possibility of 'representations and the consciousness of them' to be passed from substance to substance as motion might be passed from ball to ball is in fact doubly confused. First, the idea of thought is an idea of something which exists in such a way as not to depend on anything else for its existence, and as such cannot be compared to motion, which is clearly not an idea of something which exists in the relevantly independent way, but is something which cannot be understood apart from the idea of the substance of which it is a property – the ball, in Kant's example.[13] Second, Descartes' claim that 'thought … is nothing else but thinking substance itself' means that the idea, entertained by Kant, that on the Cartesian view 'representations and the consciousness of them' might be passed along from one substance

[12] *Ibid*., I.62–3.
[13] *Ibid*., I.61, the paragraph in which Descartes makes this clear.

to another is simply incoherent, *for the flow of 'representations and the consciousness of them' is itself substantival*.[14]

Strawson's attempt to press Kant's point even further by suggesting that acceptance of the Cartesian view would mean having to accept the possibility that each thought of mine might be the thought of a thousand souls simultaneously thinking this thought seems to add absurdity to misconception. No sense whatever can be made of the idea that my thought 'It's quite a nice day' might be a thousand such thoughts. In any case, this claim of Strawson's clearly rests on the fundamental misinterpretation of Descartes that I have just emphasised. What remains to be considered, I think, is a concern about the identity of the thinking self through time, and the idea that the Cartesian conception of a thinking substance is so peculiarly vulnerable to doubts about continuing identity as to render the whole idea unusable.

I have two points to make about this suggestion. The first is that it seems to me a mistake to suppose that doubts about the continuing identity through time of the thinking substance (*res cogitans*) are of an essentially different order from the doubts that might be raised about the continuing identity of any physical object. I can always raise a doubt about whether the object in front of me continues as the same object, or whether it is replaced by an exactly similar object each time I blink. This sort of doubt is certainly 'hyperbolic' to use Descartes' word, but it is not logically absurd. Doubts about the continuing identity of the self when one is going through a complex argument, say, or being held by a piece of music, seem to be of the same order. We ought to see that they are indeed of the same order once we free ourselves of the Kantian misconception that for Descartes thoughts inhere in something which is not itself thought, but an unknowable substratum underpinning thoughts. That conception certainly invites the sceptical response expressed by Kant, Strawson and others. There is a very great deal more that I might say about personal identity through time, but I cannot enter more deeply into that issue here.[15]

[14] After it first occurred to me that the usual interpretation of Descartes' view of substance was a travesty I discovered that Galen Strawson has made the same point in a number of places. However, he actually endorses the footnote of Kant's for what seem to me to be mistaken reasons. I have no space to pursue my disagreement with Strawson on this point and on substance dualism in general, particularly with regard to his claim that Descartes' argument for the real distinction between mind and body was an error.

[15] I have said quite a lot about this issue in *The Identity of the Self* (Edinburgh: University Press, 1981) and in a number of papers, including 'Personal Identity and the Idea of a Human Being' in D. Cockburn (ed.),

The other point that needs to be made is this. I have argued that the relation between the intentional and the physical is causal: intentional mental events are causes of behaviour, and the intentional does not reduce to the physical, nor can it be seen to be realised in the physical. Given this, we must ascribe to the mental the status of substance. To go back to Kant's example, it is certainly true that the motion of one ball might be passed on to the next in some sense, but it is not some theoretically separable property, motion, which has this causal power, the power to cause the next ball to move, but the object-in-a-state-of-motion, and that object is certainly a substance. To ascribe causal power to the mental, and to reject the idea that the mental is a property of the physical, or is reducible to the physical, or is realised in the physical, is to be committed to the view that the mental is substantival. So we have interactionist substance dualism.

6. Why Kim's Rejection of Immaterial Minds is Unsuccessful

It has been a stock objection to Cartesian dualism that the notion of a causal interaction between the material and the immaterial is unintelligible. Yet, as Kim acknowledges, it is very difficult to come up with a conclusive demonstration that the idea of trans-substantival causal transactions is incoherent, and, as he mentions, Descartes actually talks of 'an association between thoughts and bodily motions or conditions so that when the same conditions recur in the body they impel the soul to the same thought; and conversely when the same thought recurs, it disposes the body to return to the same conditions'.[16] What is notable about this passage, as Kim points out, is that Descartes actually posits, not a Humean conjunction between the material body and the immaterial mind, but the operation of causal power across the substantival divide. The challenge to Descartes' critics is, then, to show just why this is incoherent, 'to put up a real argument or shut up', as Kim puts it.

Human Beings (Cambridge: Cambridge University Press, 1991), 127–42, and 'Personal Identity and Objective Reality' in J. J. MacIntosh and H. A. Meynell (eds), *Faith, Scepticism and Personal Identity* (Calgary: University of Calgary Press, 1994), 185–98.

[16] See J. Kim, *Physicalism, or Something Near Enough*, *op. cit.*, 75, where the passage from a letter of Descartes is quoted.

Kim claims that he has such an argument.[17] 'Descartes' trouble', he claims, 'has nothing to do with the bruteness or primitiveness of causation or whether causation is merely a matter of Humean regularity, and it has everything to do with the supposed nonspatiality of Cartesian minds'. Here is the argument:

> Suppose that two persons, Smith and Jones, are 'psychologically synchronised', as it were, in such a way that each time Smith's mind wills to raise his hand, Jones' mind also wills to raise his (Jones') hand … There is a constant conjunction between Smith's mind willing to raise a hand and Smith's hand's rising, and, similarly, between Jones' mind's willing to raise a hand and Jones' hand going up … But there is a problem. For we see that instances of Smith's mind's willing to raise a hand are constantly conjoined not only with his hand's rising but *also with Jones' hand's rising*, and, similarly, instances of Jones' mind willing to raise a hand are constantly conjoined with Smith's hand rising. So why is it not the case that Smith's volition causes Jones' hand to go up, and that Jones' volition causes Smith's hand to go up?
>
> It will not do to say that after all Smith wills *his* hand to rise and that's why his willing causes his hand, not Jones' hand, to rise … The reason is that what makes Smith's hand Smith's, not Jones', that is, what makes Smith's body the body with which Smith's hand is 'united' is the fact that there is a specially intimate and direct causal commerce between the two. To say that this is the body with which this mind is united is to say that this body is the only material thing that this mind can *directly* affect … This is *my* body, and this is *my* arm because they are things I can move without moving any other body …[18]

Thus the claim that an immaterial, or non-spatial, mind causally interacts with a body in space is incoherent, Kim would have us believe. In fact the argument is quite unsuccessful. What undermines it is the claim that 'to say that this is the body with which this mind is united is to say [only] that this body is the only material thing that this mind can *directly* affect'. This claim is simply false. To see this, let us compare Kim's claim with what A. J. Ayer said about this issue in his

[17] It is only later in his discussion that Kim acknowledges that the argument (the 'pairing problem') was first presented by John Foster in 'Psychological Causal Relations', *American Philosophical Quarterly* **5** (1968), 64–70: *ibid*., 79.

[18] *Ibid*., 76–7.

paper, 'Privacy'. In that paper, Ayer lists three conditions in virtue of which some particular body is mine. The body that is mine is the body which (a) is under the control of my will in a way in which no other body is, (b) is delineated by my sensations, and (c) provides, as it were, the centre from which I view the world.[19] This seems to me exactly right. And of these three conditions, only the first is causal; that my body is the one which is the locus of my sensations, and that it provides the centre from which I view the world, are conditions which do not involve causality in any way. I therefore find no obstacle to the claim that when I will my arm to go up, it is my arm, not simply in virtue of the fact that it is a thing which I can directly affect, something which, in Kim's speculation, would also be true of an arm attached to another body, but in virtue of that arm being a part of the body which is delineated by my sensations. And it is also, crucially, that arm to which my volitional thought is directed.

You would think then that there can be nothing which might prevent Kim's acceptance of causal interaction. For surely the notion of a body's being one's own in virtue of being thus delineated cannot present a problem. But if Kim can accept that, then his fundamental claim that what makes an arm Smith's is just that there is a specially direct causal connection between Smith and that arm must be rejected. My volition is directed to the particular right hand, say, which is mine in virtue of being part of the body delineated by my sensations. What is the problem?

Two points of confusion seem to be evident in Kim's discussion of this issue. First, it is very difficult to make any sense of what Smith's willing to raise a hand could be in the example as presented by Kim. Are we to suppose that Smith's willing is not directed to one particular hand, or what? Even if we can accept Kim's speculation that Smith's willing is always followed by the hands of two different bodies rising, the question remains, which hand did Smith in fact will to rise? Was it both of them, or was his willing a sort of objectless mental operation?

The second problem for Kim is what seems to me to be his extraordinary misconception as to the nature of intentionality. Kim claims that 'we need causal relations to understand intentionality'. If this were the case, then there might indeed be the problem that Kim supposes that he has highlighted. For in Kim's imagined case there appears to be an equally direct causal connection running from Smith both to Arm A (Smith's arm, as we would want to say) and

[19] A. J. Ayer, 'Privacy' in A. J. Ayer, *The Concept of a Person and Other Essays* (London: Macmillan, 1963), 55–6.

to Arm B (Jones' arm). But the claim that intentionality rests on causality is simply false. Kim is led to make this claim by generalising from the case of perception. We can allow that what it is for me to perceive a certain tree rather than a qualitatively indistinguishable one hidden behind it is a matter of the causal impact of that tree on my visual experience. But to suppose that what is true of perception is true of intentionality in general is a clear mistake. I may, for example, be looking at two trees, but only of the one on the left do I think that it is blocking out too much light and may well have to come down. That such a thought is directed to the one tree rather than the other is not a matter of causal connection between that tree and my consciousness at all. I may be thinking, not about one of the trees in front of me, but about the Battle of Hastings or the Big Bang, and no reference to causality is required to understand these examples of intentionality. Once this (surely obvious) point is seen, that is, once it is seen that intentionality is not a matter of a causal process running from an object in the world to the subject, but of the directedness of thought from the subject to an object in the world,[20] then surely there can be no obstacle to accepting that, as I said above, I think volitionally of that right arm which is part of the body delineated by my sensations, and which provides the centre from which I view the world.

7. Why Kim's Rejection of the Possibility of Immaterial States Being Connected to the Body is not Coherent

Kim has another concern, which seems to me quite independent of the particular problem about causality which he discusses. This is the problem of how we can make sense of something which is immaterial and non-spatial being connected to or situated in some particular body in space. If the soul is non-spatial, how can it be located in a particular body, or even be connected with some particular body?

We must ask how seriously Kim can press this difficulty, if such it be, in the light of his own view of the nature of qualia or sensations. Kim admits that it is not possible to give a functionalist or physical reduction for sensations. They are themselves, then, not part of the physical world. But it would be a bold person who claimed that they have no position in space at all. On the contrary, we feel them in various areas of the body. The fact that Kim regards them as

[20] Though not necessarily to an object *in the world*, of course, since there may be no such object, as in the case of fear of ghosts.

epiphenomena, as the effects of processes in the body, does not affect the claim that they themselves are not events or objects in the physical world. The fact that they are caused by locatable events in the body is not enough to secure their immaterial effects such a location. But they are, or seem to be, located in the physical world.[21] Given this, I do not see how Kim can press the objection that if the mental is conceived of as immaterial then there is no sense in which it can be located in space.

Further than this, Kim's claim that sensations such as that of pain, which are clearly locatable in the body, are mere epiphenomena, having no effect on behaviour at all, is wholly unconvincing. Pains and itches, he says, have a motivational/behavioural aspect in addition to their qualitative aspect, and 'it is clear that the motivational/behavioural aspect of, say, pain, can be given a functional account'.[22] I think this attempt to sever the motivational aspect of pain from its qualitative aspect is totally incredible. Is it seriously claimed that a world in which the qualitative aspect of pain was totally absent, a world in which no one had ever felt pain, would have been just like this one, a world in which people over the centuries have devised hideously painful ways of putting men and women to death, such as burning at the stake and being hanged, drawn and quartered? What is the current debate on whether torture is ever admissable about, if it is not about whether it can ever be right to inflict such horribly unpleasant experiences on anyone? And in contrast with the case of pain, what can we be doing on Kim's view when we recommend a certain sensory experience – a scent, or a taste, say – as particularly pleasant? How can this be reconciled with the claim that sensations are mere epiphenomena, having no place in the causally determined order of the world? The attempt to hive off qualia as items which have no causal relevance is clearly quite misconceived. Such qualia, then, are, by Kim's own admission, not items in the physical world, but are clearly locatable, or so it seems, and clearly have effects on our behaviour.

I conclude, then, that, however much it may grate against contemporary positions and assumptions, substance dualism remains the most persuasive solution to the mind-body problem.

University of Edinburgh

[21] There are complications, of course, which relate to Descartes' description of sensations as 'confused modes of thinking'. I cannot pursue this matter here, but I do not think that it has a crucial bearing on the argument.

[22] J. Kim, *Physicalism, or Something Near Enough*, *op. cit.*, 170.

Fundamental Singleness: How to Turn the 2nd Paralogism into a Valid Argument

GALEN STRAWSON

Abstract

[1] Experience (*i.e.* conscious experience) is a real concrete phenomenon. The existence of experience entails the existence of a subject of experience. Therefore subjects of experience are concretely real (or at least one is). [2] The existence of a subject of experience in the lived present or living moment of experience, *e.g.* the period of time in which the grasping of a thought occurs, provably involves the existence of singleness or unity of an unsurpassably strong kind. The singleness or unity in question is a metaphysically real, concrete entity. So if thoughts, or any experiences at all, really do occur or exist – and they do – then there exist entities that are genuine, concrete, metaphysical unities of an unsurpassable sort. [3] There is a metaphysically irreproachable sense in which we may – must – take these unsurpassable metaphysical unities to be themselves (a) subjects of experience, although we may also take them to be (b) thoughts or experiences. If so, there is a sound argument (using Kantian materials) to the conclusion of the 2nd Paralogism. [4] Perhaps (a) and (b) are not in the final analysis distinct. Perhaps Kant is right, in his 1772 letter to Herz, that 'the thinking or the existence of the thought and the existence of my own self are one and the same'.

1. Introduction

One's experience – one's overall or total field of experience – is *complex*, considered at any given time, and there's also an extremely fundamental sense in which it's *unified*, considered at any given time. One of the many reasons why philosophers have posited an entity called 'the self' has been to furnish themselves with something that might possibly help to explain the unity of the complex at a given time. Something, it seems, must tie or bind the complex or 'manifold' of experience into a unified whole, or somehow constitute it as a unified whole. The self is posited as that which performs this operation, or is at least the locus of this unity, and is therefore posited as something that is (therefore) in some way ontologically additional to the elements of the complex.

The aim as so far described is to explain the *synchronic* unity of experience. Many also appeal to the idea of the self when giving an

doi:10.1017/S1358246110000159

account of the *diachronic* unity of experience. It seems to me, though, that Hume, Kant, William James and others have fatally undermined the idea that we need to postulate a single persisting metaphysical entity called 'the self' in order to explain the phenomena of the diachronic unity of experience (I'm not entirely sure what these phenomena are meant to be), so I'm going to restrict my attention to the synchronic case. By 'synchronic', though, I don't mean a durationless instant, because there can be no experience at all at a durationless instant. I mean experience during a short temporal interval which I'll call the 'living moment of experience', taking this term to cover such things as the time in which the grasping of a thought occurs.[1]

I'm going to argue that the self exists, and that there's a very close relation between the self and the (synchronic) unity of experience. I don't, however, think it's any sort of explanatory relation. I'm also going to argue that selves themselves are unities, in a certain strong sense, and indeed that they qualify for the title 'object' (or even 'substance') in fundamental metaphysics – if any things do. But I'm not going to give them any explanatory work to do, relative to the phenomenon of the (synchronic) unity of experience. And I'm going to talk about subjects of experience rather than selves. This shouldn't lead me astray, for if selves exist, then they're certainly subjects of experience, in which case anything I establish about subjects will also hold true of selves. That said, I'm going to focus on a restricted notion of the subject that will lead some – perhaps most – to deny that what I think of as subjects are what they think of as selves.

2. Experience

I start from the point that consciousness – conscious experience or as I will simply say *experience* – exists:

[1] experience exists (is real).

[1] More strictly, I take the living moment of experience to be a very short period of time, as short as the shortest period of time in which experience can be said to be going on at all, and therefore much shorter than the so-called 'specious present', which probably has a maximum extent of around a third of a second. (I follow Dainton in this estimate – see his *Stream of Consciousness: Unity and Continuity in Conscious Experience* (London: Routledge, 2000), 171 – although he may now be reducing it.)

This is obvious, but I'll list it as an assumption. By 'experience' (mass term) I mean the experiential-qualitative character or 'what-it's-likeness' of experience***s*** (count noun) where this experiential-qualitative character is *considered just as such*, and is therefore considered wholly independently of anything else that may be supposed to be part of what experiences consist of when they're considered as a whole; and, more generally, wholly independently of anything else that may be supposed to exist when experiences exist. The relevant conception of experience and experiential content is therefore wholly 'internalist', in present-day terms.[2]

What do I mean by 'experience'? Basic examples will do – pain, the experience or what-it's-likeness of seeing the colour red, of tasting bananas, of finding something funny. I mean things like these. (I'll continue to append 'or what-it's-likeness' to 'experience' for the moment, although I've defined them as synonymous.) What is it to be a *real realist* about experience, *i.e.* about what-it's-likeness (pain, itching, the experience of seeing the colour red or of tasting bananas)? The best reply, perhaps, is that to be a real realist about experience is to continue to take colour experience/what-it's-likeness or taste experience/what-it's-likeness or pain experience/what-it's-likeness to be what one took it to be – knew it to be – quite unreflectively, before one did any philosophy, *e.g.* when one was five.[3] However many new and surprising facts real realists about experience learn from scientists – facts about the neurophysiology of experience, say, or the 'filling-in' that characterizes visual experience, or 'change blindness', or 'inattentional blindness'[4] – their basic grasp of what colour experience or pain experience (*etc.*) is remains *exactly the same as it was before they did any philosophy*. It remains, in other words, correct, grounded in the fact that to have experience at all is already to know what experience is, however little one reflects about it. When it comes to experience, 'the having is the knowing'.

[2] I take 'what-it's-likeness' to cover the total character of our experience, everything that the Phenomenologists take as their object of study, and to be as much cognitive as sensory.

[3] One doesn't need to attribute a general conception of experience to five-year olds to make the point, but one shouldn't underestimate five-year-olds – or think that they can't make a distinction between the character of their experience = what-it's-likeness and what it is experience of.

[4] See *e.g.* L. Pessoa and P. De Weerd, *Filling-In: From Perceptual Completion to Cortical Reorganization* (Oxford: Oxford University Press, 2003); D. J. Simons and D. T. Levin, 'Change blindness', *Trends in Cognitive Sciences* **1** (1997), 261–7; M. Chun and R. Marois, 'The dark side of visual attention', *Current Opinion in Neurobiology* **12** (2002), 184–9.

It's knowledge 'by acquaintance'. To taste pineapple, in Locke's old example, is sufficient, as well as necessary, for knowing what it's like to taste pineapple.[5]

This way of saying what I mean by 'experience', and hence by 'realist about experience', guarantees that anyone who claims not to know what I mean is being disingenuous. The last paragraph wouldn't be necessary if philosophers like Dennett hadn't looking-glassed the term 'consciousness' or 'experience'.[6]

It follows from [1] that

[4] subjects exist

or that at least one subject of experience exists, because it's a necessary truth that

[2] [experience → subject]:

'an experience is impossible without an experiencer'.[7] It's 'an obvious conceptual truth that an experiencing is necessarily an experiencing by a subject of experience, and involves that subject as intimately as a branch-bending involves a branch'.[8] In other words: there can't be experience without a subject of experience, because

[3] experience is necessarily experience *for*.

Spelt out a little: experience is a matter of experiential 'what-it-is-likeness' and experiential what-it-is-likeness is necessarily what-it-is-likeness *for* someone or something. *Whatever* the metaphysical nature or category of this experiencing something, its real existence can't be denied. This point is secure even if individual-substance-suggesting noun phrases like 'an experiencer' or 'a subject of experience' are somehow misleading.[9]

I take it that [3] is a sufficient answer to those who seek to deny [2] (to claim that there can be experience without a subject), for no one

[5] I put aside the point that different people may have different taste-experiences in tasting pineapple.

[6] To looking-glass a term is to define it in such a way that whatever one means by it, it excludes what the term means.

[7] G. Frege, 'The Thought, A Logical Inquiry' in *Philosophical Logic*, edited by P. F. Strawson (Oxford: Oxford University Press, 1918/1967), 27.

[8] S. Shoemaker, 'Introspection and the Self' in *The First-Person Perspective and Other Essays* (Cambridge: Cambridge University Press, 1986/1996), 10.

[9] Descartes makes it in his 'Second Meditation' in *The Philosophical Writings of Descartes*, Volume 2, translated by J. Cottingham et al. (Cambridge: Cambridge University Press, 1641/1985), 18.

can sensibly deny [3], the essential *for*-someone-or-something-ness of experience, and I understand the notion of the subject of experience in such a way that to say that [2] experience entails a subject of experience is simply to express [3] in a different way. One could put the point paradoxically by saying that if *per impossibile* there could be pain-experience without a subject of experience, mere experience without an experiencer, contrary to [2], then there'd be absolutely no point in stopping it, because no one would be suffering.[10]

Another way to put the point is to say that

[5] [subjectivity → subject].

It's worth stating this explicitly because some like to deny [5], claiming precisely that although there obviously can't be experience without subjectivity

[6] [experience → subject*ivity*]

still there can be subjectivity without a subject – so that [2] isn't true after all.

I sympathize with the impulse behind this claim, but what it shows me is that those who endorse it are making some metaphysical assumption about what subjects are that I'm not making. Perhaps they're thinking that subjects must be things that persist for an appreciable length of time, or substances in some traditional, heavy sense of the word 'substance'. My notion of what a subject is involves no such metaphysical commitments.[11] In fact it's no more metaphysically committed than their notion of subjectivity. It rests on no more than [3], and for present purposes I can indicate its maximally non-committal metaphysical nature by saying that it has the consequence that [5] is necessarily true given that [3] is, so that [2] is secure. Given what I understand by 'subject', subjectivity entails a subject.

3. Materialism (Um-ism)

I'm going to restrict my attention to concrete phenomena like chairs, pains, and electrons, as opposed to abstract phenomena, like the idea of

[10] 'Of course there'd be a point: the universe would be a better place, other things being equal, if there were fewer pain sensations. Pains are intrinsically bad, even without owners.' Reply: Pains without owners, pains without someone-or-something pained – call them 'pains*' – are pains without pain! Pains* are impossible (apart from not being intrinsically bad).

[11] Again I follow Descartes in his 'Second Meditation'. All he claims to know at this stage of his argument is that he is *something* – not nothing – a really existing phenomenon, whatever its ontological category.

justice or the number 2 (I'm using 'phenomenon' as a general term for any sort of existent or candidate existent, suppressing its old meaning of 'appearance'). I'm going to assume the truth of *concretism*, in fact – that

[7] [real → concrete].

By 'concrete phenomena' I mean phenomena that are actually located somewhere in the universe. I take it for the purposes of argument that all location is at least temporal, if not also spatiotemporal, while noting that we may be very wrong about the nature of such location.[12] Immaterial souls, then, and the vaguest fleeting feelings, are 100% concrete. They may seem flimsy compared with a bullet, but that's a superficial view of things. It has nothing to do with metaphysical concreteness.

I'm going to assume that monism and in particular materialism is true, *i.e.* that all concretely existing phenomena are physical phenomena:

[8] [concrete → physical][13]

which with [7] entails

[9] [real → physical].[14]

The converse of [8], *i.e.*

[10] [physical → concrete]

is obviously true, and [8] and [10] sum to

[11] [concrete ↔ physical]

from which I'll take it to follow that

[12] [concrete object ↔ physical object].

[12] I take 'temporal' and 'spatiotemporal' to be referring terms that pick out an objectively real 'dimensionality' (the dimensionality of the concrete real) however wrong we are about the nature of this dimensionality in so far as we take ourselves to know anything more about this nature than what is given in a mathematical characterization of it (*sc* of the set of homomorphic dimensionalities of which it's a member).

[13] [8] on its own amounts to materialism as the term 'materialism' is used in the philosophy of mind (to rule out things like immaterial souls); but materialism can also be understood in a stronger sense which incorporates [7], the assumption that all real phenomena are concrete phenomena.

[14] Some reject [8]. They say that there are 'abstract objects', numbers for example, that are part of reality in every sense in which concrete physical objects are ('abstract' and 'concrete' are defined as mutually exclusive and exhaustive opposites).

In this discussion, then, one can always replace 'concrete' by 'physical', and conversely.

[1] and [9] entail

[13] [experience → physical],

and I'm not just a materialist, I'm also, given [1], a 'real' materialist, a realistic materialist, a philosophically serious materialist, *i.e.* a materialist who is fully realist about the existence of (conscious) experience. Real materialism holds that experience – the concretely existing phenomenon of 'what-it's-likeness' considered just as such and so wholly independently of anything else – is wholly physical, although physic***s*** has no resources for characterizing its nature so considered.[15]

Having declared my materialism, I should say that I take 'material' and 'physical' to be *natural-kind* terms; and natural-kind terms are paradigmatic examples of terms which succeed in referring to what they're introduced to refer to even though we who introduce and use them may be very wrong about the nature of what they refer to. It's plain, I think, that there's really no less reason for us to take 'material' and 'physical' to be natural-kind terms than there is for us to take terms like 'gold' and 'tiger' to be natural-kind terms. But although almost all present-day philosophers take 'gold' and 'tiger' to be natural-kind terms, very few do so in the case of 'physical' or 'material'.

One reason for this, perhaps, is that we find it hard to take 'temporal' and 'spatiotemporal' to be natural-kind terms, in the way I endorsed above; and one thing then leads to another. Even if this isn't so, it's easy to see why there's so much resistance to taking 'physical' to be a natural-kind term. To do so is to take it to be a term that certainly refers to the reality that physics concerns itself with while allowing that we (and physics) may be very wrong about fundamental aspects of the nature of this reality. This, however, is to allow that the ultimate nature of the thing that physics is about, the physical as denoted by the natural-kind term 'physical', may possibly be irreducibly mental or experiential in character: not only in part, as I've already proposed, but even perhaps as a whole. Eddington and Russell are among those who suppose that the physical (natural-kind-term use) is or may be wholly mental or experiential in character, and I'm inclined to agree with them, if only for reasons of theoretical parsimony. Few, though, can stomach this consequence of taking 'physical' as a natural-kind term. They insist

[15] See *e.g.* G. Strawson, 'Real materialism' in *Real Materialism and Other Essays* (Oxford: Oxford University Press, 2003/2008).

on its retaining a descriptive force that goes beyond anything warranted by physics.[16]

This is understandable. To take 'physical' to be a natural-kind term amounts to taking it to be fully semantically equivalent to the neutral term 'real and concrete', so that the difference between materialism and mentalism or idealism, say, can no longer be expressed. Common terminology is further offended when it's proposed – by Russell among others – that we can know the non-structural or non-mathematically-characterizable nature of the physical only in being directly acquainted with the character of our experience (the having is the knowing). I continue to call myself a materialist in order to stress the point that my view is [i] a view about the actual nature of the concrete reality that physics deals with and [ii] a view that allows (takes it) that physics may have got a very great deal right about the (ultimate) nature of reality, but I know that this natural-kind term use of 'physical' isn't going to catch on. It's going to continue to be judged to be absurdly eccentric, and indeed a misuse of the word. Let me therefore introduce the term 'um' as a synonym for the natural-kind term use of 'physical'. I'm an um-ist (an *Ur*-materialist). But I'm going to continue to use the term 'physical' for the most part, because the materialist or um-ist claim that matters most at present is simply the claim that experience is wholly a matter of brain activity, and until the word 'um' is widely used (surely it will not be long), the fact that I take physics and neurophysiology to have got a very great deal right about the brain is better kept in focus by using the word 'physical' than the word 'um'.

So far, perhaps, so good. But I've spoken of objects, and said that I'm going to argue that subjects of experience are objects –

[20] [subject (of experience)→ object]

– so I need to say what an object (or 'substance') is. This is a very big, very general and very difficult question in metaphysics. I think we need to renounce tables and chairs as paradigm cases of objects in order to approach it (which is not to say that tables and chairs won't in the end qualify as objects).

4. Object [1]

The central thing to say, I think, is that whatever an object is, it is – I take it – a certain sort of *unity*

[16] On this 'structural realist' point see *e.g.* G. Strawson, 'Real Materialism' *op. cit.* and, for a sustained treatment, J. Ladyman and D. Ross (with D. Spurrett and J. Collier), *Every Thing Must Go: Metaphysics Naturalized* (Oxford: Clarendon Press, 2007).

[14] [object → unity (of some sort)]

and if it's a concrete object then it is – I take it – a certain sort of concrete unity

[15] [concrete object → concrete unity]

and if it's a physical object, as it is if materialism is true, then it is – I take it – a certain sort of physical unity

[16] [physical object → physical unity].

Since I'm restricting my attention to the concrete (and hence the physical), I'll regularly drop the words 'concrete' and 'physical' from now on.

What kind of unity is in question? I'm going to propose that if anything is going to qualify as an object in fundamental metaphysics (by which I mean – at least – metaphysics that gives no special weight to the ordinary categories of thought), then it must be a *strong* unity

[17] [object → strong unity]

in a sense I'll try to illustrate. I think, in fact, that strong unity is not only necessary but also sufficient for objecthood

[18] [strong unity → object]

which sums with [17] to give

[19] [strong unity ↔ object]

the view that unity, of a certain sort which remains to be specified, is the sole and sufficient criterion of objecthood.

I'll go on to claim that subjects of experience qualify as strong unities, and that there are in fact no better candidates for the title 'strong unity' than subjects of experience

*[20] [subject → strong unity].[17]

I've asterisked [20] to mark the fact that it's something I want to try to explain or argue for, *i.e.* not something I'm [a] taking to be evident, or [b] taking to be true by definition, or [c] assuming without significant discussion, or [d] claiming to be able to derive from things in categories [a]-[c]. I can derive

[21] [subject → object]

from *[20] if I can make a sufficient case for [18].

[17] Note that Cartesian souls are unbeatable examples of concrete strong unities, even if they're ruled out by materialism.

Clearly I need to say more about the notion of a strong unity. Before that, though, I should say more about what I mean by 'subject of experience'.

5. Subject

There are various different conceptions of what a subject of experience is. I'll consider three and focus on one. There is, first,

[C1] the *thick* conception according to which human beings and other animals considered as a whole are subjects of experience.

Secondly, there is

[C2] the *traditional inner* conception of the subject according to which a subject of experience properly or strictly speaking is some sort of persisting inner locus of consciousness, an inner someone, an inner mental presence.[18]

Both [C1] and [C2] build in the natural assumption that a subject may and standardly does continue to exist even when it's not having any experience. For whether you think that human subjects of experience are whole human beings, or whether you think they're inner loci of consciousness, you're likely to allow that they can continue to exist during periods of complete experiencelessness, *e.g.* periods of dreamless sleep.

In this respect [C1] and [C2] contrast with a third conception of the subject

[C3] the *thin* conception of the subject according to which a subject of experience is an inner thing of some sort that exists if *and only if* experience exists of which it is the subject.

The thin conception of the subject stands opposed to both [C1] and [C2] precisely because they both assume something that [C3] rules out: the idea that a subject of experience can be said to exist in the absence of experience. As a materialist (um-ist) I take it that the goings on that wholly constitute the existence and experience of a thin subject consist entirely of activity in the brain.[19] A thin

[18] As a materialist I understand 'inner' literally, in a straightforwardly spatial sense.

[19] This claim isn't 'reductive' or experience-denying in any way, given real um-ism – one might call it 'adductive'. It doesn't claim that experience is anything less than what we ordinarily conceive it to be, but rather that the brain is more than what many people ordinarily conceive it to be.

subject, I take it, is a certain sort of neural 'synergy', a neural synergy that is literally (spatiotemporally) part of the neural synergy that is an occurrence of experience – experience which is of course necessarily-subject-of-experience-involving experience.[20]

I face a problem of exposition, because most are so accustomed to [C1] and [C2], and to the idea that [C1] and [C2] exhaust the options, that they can't take [C3] seriously. And yet [C3] simply makes a place for a natural use of the term 'subject' according to which it's a necessary truth, no less, that there can't actually be a *subject of experience*, at any given time, unless some *experience* exists for it to be a *subject of*, at that time

i.e.

[22] [subject of experience → experience]

which couples with [2] to give

[23] [experience ↔ subject of experience].

Given the thin conception of the subject, there can no more be a subject without an experience than there can be a surface without extension. The thin conception requires that the subject be *live*, as it were, in order to exist at all. There is no dispositional use of the expression 'subject of experience'.

I think it's very important to have the thin or live conception of the subject available and in play in metaphysics, and from now on I'm only going to consider the thin notion when arguing for the exemplary unity of the subject. If you think the thin conception of the subject or self is eccentric, bear in mind that it's endorsed by Descartes, Leibniz, Kant, Fichte, Nozick, and many others. Kant writes, strikingly, that 'the thinking or the existence of the thought and the existence of my own self are one and the same'.[21] Hume invokes thin subjects when characterizing the notion of the subject of experience that is philosophically legitimate given his strict empiricist principles (*i.e.* when expounding the 'bundle' theory of the self): 'when my perceptions [experiences] are remov'd for any time, as by

[20] Elsewhere I consider the suggestion that the subject synergy is not just part of the experience synergy, but identical with it (G. Strawson, 'What is the relation between an experience, the subject of the experience, and the content of the experience?' in *Real Materialism and Other Essays*, *op. cit*).

[21] I. Kant, Letter to Marcus Herz, February 21, 1772, in *Kant: Philosophical Correspondence 1759–99*, edited and translated by Arnulf Zweig (Chicago: University of Chicago Press, 1772/1967), 75.

sound sleep', he writes, 'so long am I insensible of *myself*, and may truly be said not to exist'.[22] Sprigge writes that 'each of us, as we are at any one moment, is most essentially a momentary centre of experience or state of consciousness with the duration of the specious present'.[23]

To focus on the thin conception of the subject is not to reject the thick and traditional inner conceptions; they can all coexist. The current dominance of the thick and traditional inner conceptions does however make it important to stress that *thin subjects certainly exist as defined* – whatever the best view of their ultimate ontological category. They certainly exist because experience certainly exists, and to speak of a thin subject is just to speak in a certain way of a feature of reality that certainly exists given that experience exists ([2]). In fact it's to speak of a feature of reality that's an essential *part* of what it is for experience to exist. One can put the point by saying that experience necessarily involves experienc*ing* (experience just is experiencing) and the existence of a thin subject is guaranteed by the fact that there's experiencing.

I suspect that the human process of experience is as a matter of empirical fact constantly interrupted, and that thin subjects are consequently short-lived in the human case, but ephemerality isn't an essential part of their definition. To believe that experience is unbroken throughout any period of wakefulness in a normal human being is to leave room for the view that thin subjects can last at least as long as human beings can stay awake. To believe that experience never ceases in a normal human being, from the first moment of experiential

[22] *A Treatise of Human Nature*, edited by D. F. Norton and M. Norton (Oxford: Clarendon Press, 1739–40/2000), 165. Contrary to widespread belief, Hume never claims that there's no subject or self at all when there's experience. The point is sufficiently established by the quotation, but it would be no less certain without it. It is after all a necessary truth that there can't be experience without a subject of experience, an experiencer, simply in so far as experience is essentially experience-for-someone-or-something, and Hume doesn't go in for denying necessary truths. What he denies is something much more specific: the idea that we have any good evidence for the existence of a persisting self or subject, let alone an invariable, unchanging self or subject of the sort whose existence is taken for granted by almost all philosophers in his time ('Some philosophers … imagine we are … intimately conscious of what we call our SELF: that we feel its existence and its continuance in existence; and are certain both of its perfect identity and simplicity', but we do not 'have … any idea of *self*, *after the manner it is here explain'd*' (*Treatise*, 164; my emphasis).

[23] T. L. S. Sprigge, *The God of Metaphysics* (Oxford: Oxford University Press, 2006), 474.

quickening to its final extinction, is to leave room for the view that thin subjects can last a lifetime. If we use the word 'inner' loosely to mean something like 'not identical with/not the same thing as the human being considered as a whole', and allow it to cover immaterial souls, then a Cartesian immaterial soul qualifies as a thin subject, and we make room for the view that thin subjects can last for ever.

6. Object [2]

So much for the notion of a thin subject. Now I need to go back to objects. I've proposed that there are no better candidates for being concrete objects than subjects, by which I now mean *thin* subjects. But are there any concrete objects? Is there any place for the notion of a concrete object or substance in fundamental metaphysics – in which I include physics?[24]

Some say No, *e.g.* proponents of radical 'structural realism' like the philosophers of science Ladyman, Ross, Spurrett and Collier in their recent book *Every Thing Must Go*. I take it that the idea is also rejected in Whitehead's 'process philosophy'. It's an old idea, at least as old as Nāgārjuna's *Fundamental Verses on the Middle Way*, written about 1800 years ago. I am nevertheless going to take it to be a truth of fundamental metaphysics that

[24] there is at least one object

or substance, and hence that the concept of an object or substance does have legitimate application in fundamental metaphysics.

Is it also true that the concept of an object or substance has *plural* application in fundamental metaphysics? This is a considerably more difficult question. My inclination is to agree with Spinoza that there is properly speaking only one object or substance.[25] For the purposes of

[24] See also J. Schaffer, 'Monism, The Priority of the Whole', *Philosophical Review* **119** (2010), 31–76. I take 'substance' to be a general term for a fundamental existent that has no connection to its etymological meaning of 'standing under'.

[25] The universe, or, in his terms, 'God or nature'. Parmenides agrees. Descartes agrees that there is strictly speaking only one physical object or substance (but takes it that there is a true plurality of non-physical concrete objects or substances – souls; there is also considerable support for this view in physics and cosmology). One recent version of thing-monism has it that spacetime itself is best thought of as a concrete object, a substance (not a mere dimensionality, as it were), and indeed as the only object there is. The fundamental entities currently recognized – leptons and quarks – are not strictly speaking fundamental and are to be 'explained as various

argument, however, I'm going to assume that there are many. In particular, I'm going to make the 'smallist'[26] assumption that there is a plurality of fundamental physical entities (leptons and quarks, say, or 'fields', 'simples', 'preons', 'loops', 'field quanta', …) or as I will say 'ultimates'

[25] there is a plurality of ultimates

and that they can qualify as strong unities if any things do

[26] [ultimate → strong unity]

and can therefore (given [19]) qualify as objects if any things do (and in spite of the phenomena of entanglement)

[27] [ultimate → object]

from which it follows that

[28] there is a plurality of objects.

I don't think [28] is obvious, when it comes to fundamental metaphysics, at least not if it is considered independently of the view that

[29] there is a plurality of subjects

– which might be held to be obvious, although it too is questionable, and which, with [20], entails [28].[27] The truth or otherwise of [28] is a serious question in metaphysics (as indeed is the question whether there are any objects at all). I am, however, going to assume [28] for argument.

---

modes of vibration of tiny one-dimensional rips in spacetime known as strings' (S. Weinberg, 'Before the Big Bang' in *The New York Review of Books* **44**/10 (1997)). On this view all the things we ordinarily think of as physical objects are made of rips in spacetime, the only concrete object there is. See also G. Strawson, *Selves: An Essay in Revisionary Metaphysics* (Oxford: Oxford University Press, 2009); J. Schaffer, 'Spacetime, the one substance', *Philosophical Studies* **145** (2009), 131–48.

[26] See R. Wilson, *Boundaries of the Mind: The Individual in the Fragile Sciences* (Cambridge: Cambridge University Press, 2004); S. Coleman, 'Being Realistic, Why Physicalism May Entail Panexperientialism', *Journal of Consciousness Studies* **13** (2006), 40–52.

[27] [20] depends on [18] and [19], for which some sort of case still has to be made. Note that [28] can be argued for quite independently of any theory about ultimates.

7. Interim Summary

If you've followed this far you'll know I have to do three things. I have to try to explain further what I mean by saying [17] that objects are strong unities

[17] [object → strong unity]

and equally what I have in mind in saying that all strong unities are objects

[18] [strong unity → object].

Then I have to show that (thin) subjects are strong unities

*[20] [subject → strong unity].

This will allow me to conclude that all (thin) subjects are objects

[21] [subject → object].

Might it be true not only that all (thin) subjects are objects but also that all (genuine, strong-unity) objects are subjects

[30] [object → subject]?

With Leibniz, I suspect that [30] is also true, at least in our universe, and hence, given [20], that

[31] [subject ↔ object]

is also true in our universe. This, though, is a form of panpsychism – something for another time.

One possibility at this point is to rest the idea of an object and present the argument just in terms of the idea of a (strong) unity – an idea which is perhaps clearer than, and no doubt more fundamental than, the idea of an object. I'll continue to use 'object', but one can if one wishes hear 'object' as 'strong unity', at the cost of putting up with a few tautologies ('all objects are strong unities').

8. Object [3]

The next question is this. What are the best candidates for being concrete objects, given that we want to retain the category *object* (or *individual substance*) in our fundamental ontology or metaphysics, and are committed to the view that there is more than one concrete object?

I'm going to take it in a conventional materialist way that every candidate for being an object is either an ultimate or is made up of some

number of ultimates in a certain physical relation.[28] A physical object, then, is either a single ultimate or a plurality of ultimates in a certain relation. Given that all single ultimates are physical objects, the remaining question is which pluralities of ultimates – if any – are physical objects? What should be our criterion?[29]

As before, the first and most certain thing to say about a physical object is that it is at least, and first and foremost, and essentially, *some kind of unity or singularity*. So the question is: given that ultimates themselves constitute unities of the right kind, which pluralities of ultimates do?

The phrase 'some kind of unity' is as vague as it is crucial, and some philosophers – the *subjectivists* – think that there are no metaphysical facts of the matter about which phenomena are objects and which aren't. On their view, whenever we judge something to be an object we implicitly or explicitly endorse an ultimately subjective principle of counting or individuation relative to which the phenomenon *counts as* a (single) object. I'll put this by saying that we endorse an ultimately subjective *principle of objectual unity*.

Can this be right? Many judgements of objecthood – many principles of objectual unity – are so natural for us that the idea they are in any sense subjective seems preposterous (nearly all of us think that cups, marbles, meerkats, jellyfish, fingers, houses, planets and molecules are individual objects, and there are clear pragmatic and evolutionary reasons why this is so). The subjectivists, however, are unimpressed. It doesn't follow from the fact that some judgements of objecthood are very natural for human beings, they say, that those judgements are objectively correct, or record metaphysical facts. If we were electron-sized, they say, our natural judgement about a stone might be that it was a mere collection of things, not itself a single object in any interesting sense.

It seems odd at first to think that merely subjective principles of objectual unity underlie our judgements that chairs and stones are objects, but it becomes increasingly natural as one moves away from such central cases. Although nearly everyone thinks a chair is a single object, not everyone does.[30] Although many think cities, newspapers, galaxies, and blenders (assembled from parts) can

[28] I take it that 'virtual' ultimates (virtual particles) and 'anti-matter' ultimates (anti-matter particles) are also objects, given that there are objects.

[29] This is Peter van Inwagen's 'Special Composition' question in his book *Material Beings* (Ithaca, NY: Cornell University Press, 1990).

[30] P. van Inwagen does not, in *Material Beings*, and his reasons are of considerable interest. The same goes for Nāgārjuna, in *The Fundamental*

correctly be said to be single things, quite a few do not. Although some think a body of gas is an object, many do not.

Very few think that three spoons, one in Hong Kong, one in Athens, and one in Birmingham, constitute a single thing, and yet some philosophers believe that the three spoons have as good a claim to be considered an individual object as anything else: according to one form of *universalism*, any plurality of ultimates in the universe, however scattered, counts as a single object in every sense in which a table does. A lepton in your amygdala, a quark in my left hand, and the ultimates that make up the rings of Saturn jointly constitute a single object just as surely as your pen or pet duck. No plurality of ultimates has a better claim to be an object than any other.

On the face of it, universalism is a wholly *objectivist* theory of objects. It endorses a principle of objectual unity that delivers a clear principle of counting. It tells you that if there are n ultimates in the universe then there are exactly $(2^n–1)$ objects in the universe. But it also has a highly subjectivist or 'post-modern' aura – it tells you that anything goes and everybody wins, that there's no real issue about whether any particular plurality of ultimates is an object or not – and it's arguable that genuinely objectivist positions emerge clearly only when more specific and limited principles of objectual unity are endorsed, *e.g.* by common sense, which rules in favour of tables and chairs and against the three spoons, or by Spinoza, who holds that there is only one thing or substance, the universe ('God or Nature'), or by van Inwagen, who argues forcefully that only individual ultimates and living beings – and not, say, tables and chairs – qualify as physical objects in fundamental metaphysics.[31]

I'm going to take it that there are at the very least various grades and types of physical unity, and that some candidates for objecthood have, objectively, a much better claim than others – given that we've committed ourselves to working with the notion of an object in our basic metaphysics. I think a human being has a better claim than your lepton + my quark + the rings of Saturn or the three spoons, given this commitment. There are vast numbers of merely 'conventional'

Wisdom of the Middle Way, translated with a commentary by Jay Garfield (Albany, NY: SUNY Press, c150/1995).

[31] For van Inwagen, as for Aristotle (see *e.g. Metaphysics Z* 7.1032a19, 8.1034a4), animals are the paradigm substances. I am not sure what they would make of the fact that ninety per cent of the cells in one's body are microbial cells.

unities or objects (to use Buddhist terminology) that are not of central concern in fundamental metaphysics, but there are also perhaps, certain irreducibly real objective unities, or objects.[32]

With this in place, consider the following suggestion. As one advances in genuine um-ism, deepening one's intuitive grasp of the idea that *experiential*, *mental* phenomena are physical phenomena in every sense in which non-experiential, non-mental phenomena are,[33] one of the things that becomes apparent is this: that when it comes to deciding which phenomena in the universe count as objects and which do not, there are *no* good grounds for thinking that *non-experiential, non-mental criteria or principles of unity* – of the sort that we use to pick out a dog or a chair – are more valid than *mental or experiential criteria or principles of unity*. It's arguable, in fact, that there is no more indisputable unity in nature, and therefore no more indisputable physical unity or singularity, and therefore no better candidate for the title 'physical object', than the unity that we come upon when we consider the phenomenon of the (thin) subject of experience as it exists in the living moment of experience, experiencing seeing books and chairs and seeing them as such, say, or consciously comprehending the thought that water is wet – an event that necessarily involves the concretely occurring thought-elements WATER and WET forming a true unity of some sort, a unity without which the thought *water is wet* can't be said to have occurred at all.[34] I can think of no better candidates in concrete reality for what I'm calling 'strong' unity, unity that is not only [17] necessary for genuine objecthood (if there is such a thing) but also – I am going to propose – [19] sufficient. As far as I can see, the only serious (and mutually excluding) competitors for equal first place are

(a) the universe,

to be identified, perhaps, with spacetime considered as a substance, and, lagging somewhat behind,

(b) individual ultimates

[32] For further discussion see *e.g.* P. van Inwagen, *Material Beings*, *op. cit.*

[33] Assuming that there are any non-experiential phenomena (it's a fundamental feature of real materialism that 'physical' doesn't entail 'non-experiential').

[34] There's no implication here that the subject of experience is or must be thought of as an *agent* that *brings about* the binding or seizing, and this isn't I think a helpful or necessary idea.

– if indeed there are any. Between these two extremes, and given the absolute centrality of unity or singularity considerations when it comes to determining objecthood, it's arguable that subjects of experience as just characterized are the best qualified plurality-of-ultimates-involving candidates there are for the status of physical objects.[35]

9. Fundamental Singleness

That's the idea, at any rate. The idea is to show that subjects are strong unities and therefore objects, and the subjects I have in mind are 'thin' subjects as they exist in the living moment of experience. Some may think that thin subjects in the living moment of

[35] I like the arguments of the physicist Richard Feynman and the philosopher Peter van Inwagen that things like chairs are really pretty inferior candidates for being objects, when one gets metaphysically serious; but I'm not appealing to them here. Note that to claim that there are no better candidates for the title 'physical object' than the unities we come upon when we consider subjects of experience in the living moment of experience isn't to claim that we come upon living-moment-of-experience-sized objects. If time is dense and periods of experience are continuous, then living moments of experience are theoretical abstractions from a continuum, not genuinely discrete entities. In this case the qualifying phrase 'considered in the living moment of experience' doesn't chop the subject at the boundaries of the living moment of experience (which don't exist) to deliver a distinct living-moment-of-experience-sized object. (The objects in question must be theoretical abstractions if time is dense because although they're countable in the mathematical sense – countably infinite – and although *countable* suggests *discrete*, no concrete infinity of things can exist in any finite period of time. To think that there could be a concrete infinity in a finite period of time is, in effect, to think that infinity is or could be finite.) It's an empirical question how long thin subjects last, and there's no good reason to suppose that the analytical cut that thought makes in considering the subject in the living moment of experience corresponds to a real division in nature. Questions about the temporal extent of objects must in other words be a matter of natural fact, and can't depend on what we can intelligibly isolate as objects of thought. If time is real but is not a continuum, consisting of discrete 'chronons', as some suspect, then living moments of experience may be chronon-sized, which is nice and simple; and there are other ways in which things could turn out to be simple – if, for example, experience-constituting activity in the brain comes in discrete pulses, 40 a second, say, or perhaps 1,000. The 'temporal parts' that analytic metaphysicians play with are (as usual) of no ontological interest whatever.

experience are intolerably odd candidates for being objects. The strength of their candidacy follows from the idea that strong-unity considerations are paramount when it comes to establishing claims to objecthood in fundamental metaphysics. 'No entity without identity', as Quine used to say: no object without 'identity conditions', 'criteria of identity'; and no concrete object, presumably, or at least no spatiotemporal object, without synchronic and diachronic identity conditions. In the synchronic case, one needs to be able to draw a line round the object – here a thin subject – in thought in such a way as to be able to say, at a given moment t_1, 'Here's one thin subject, and here's another'. In the diachronic case, considering thin subjects as things that may last longer than the living moment of experience, and restricting our attention to single brains for simplicity, we presumably need to be able to draw a line round them in such a way as to be able to say 'This is one thin subject, existing from time t_1 to time t_2, and this, now, is another different one, existing from t_2 to t_3, or from t_3 to t_4'.

I say 'presumably' because I'm not sure that we do have to be able to give an account of the diachronic identity conditions of something – even in principle – in order to have reason to say that that something is an object or substance (given that we want to retain the categories of object and substance in our metaphysics). I'm not sure what I would say to someone who insisted that I could no longer claim to be dealing with the notions of object and substance at all, in abandoning the requirement of specifiable diachronic-identity conditions, but I'm going to put this question aside and restrict attention to the synchronic case.[36]

Consider, then, an experiencing subject at a given moment, and consider the totality of its experience at that moment. Consider its 'total experiential field', the total content[37] of its experiential field at a given moment, and ask whether it could fail to be experientially single or unified in being what it is.

To ask the question is to see the fundamental sense in which the answer is No. The total experiential field involves many things – rich interoceptive (somatosensory) and exteroceptive sensation, mood-and-affect-tone, deep conceptual animation, and so on. It has, standardly, a particular focus, and more or less dim peripheral areas, and it is, overall, quite extraordinarily complex in content. But it is for all that a unity, and essentially so. It is fundamentally unified, utterly indivisible, as the particular concrete phenomenon

[36] See G. Strawson, *Selves*, *op. cit.*, 403.
[37] Content is internalistically understood as before.

it is, simply in being, indeed, *a* total experiential field; or, equivalently, simply in being *the* content of the experience of a single subject at that moment.[38] The unity or singleness of the (thin) subject of the total experiential field in the living moment of experience and the unity or singleness of the total experiential field are aspects of the same thing. They're necessarily dependent on each other even if they aren't the same thing.

I think, in fact, that there's only a conceptual distinction between them, not a real distinction, in Descartes' sense, for neither can conceivably exist without the other. In which case there's no dependence relation of a sort that requires two really (real-ly) distinct entities.

The present point, however, is independent of any dramatic metaphysical claim about the ultimate metaphysical identity of the unity of the (thin-subject) experiencer and the unity of the experiential field (or the experiential content, or, most simply, the experience). It's simply the point – the simple point – that *a* field, a single field, entails *a* subject, a single subject, and conversely. If you bring a single thin subject into existence at t_1 then you necessarily bring a single, unified experiential field into existence at t_1, and conversely. The singleness of the (thin) subject at t_1 entails, and is essentially constitutive of, the singleness of its total experiential field at t_1. So too, the singleness of the total experiential field at t_1 entails the singleness of its subject at t_1,[39] for the material that is experienced is necessarily experienced in a single or unified experiential perspective.[40] The point is a 'logical' one, as 'trivial' as it is important.

I'll call it *Fundamental Singleness*. It applies to all necessarily-subject-of-experience-involving experiential-field unities, whatever

[38] There's no need to rule out the possibility that there can be more than one total experiential field in a human being at a given time. If there can, the present point about necessary singleness will apply to each one individually (see G. Strawson *Selves*, *op. cit.*, 2.19–2.20).

[39] The former is also arguably essentially constitutive of the latter – in which case the asymmetry of the 'X constitutes Y' relation lapses, and simple identity is indicated – but I won't press the point. Elsewhere I consider the idea that the experiential content *c* of the experience of the subject *s* is related to *s* 'in such a way that its being is at least partly *constitutive* of the being of *s*. On this view *c* is, as it were, the body or flesh of *s*, without which *s* … cannot exist and is nothing. *s*, we feel, cannot simply be the same as *c*, but *s* is nothing without *c* – not just utterly empty, but non-existent' (cf. G. Strawson, *Selves*, *op. cit.*, 410).

[40] Anything that is experientially there in any sense at all is automatically part of the unity – just as any paint intentionally added to a painting by the painter is automatically part of the picture.

their duration, and to all possible subjects of experience, however primitive. To this extent it's independent of Kant's much higher-level claim that when we consider ourselves as subjects we must necessarily apprehend ourselves as single; his claim (in the Paralogisms section of the *Critique of Pure Reason*, which I will call 'the Paralogisms') that 'the proposition *I am simple* [*i.e.* single] must be regarded as a direct expression of apperception'.[41] For one thing, his claim is about expressly self-conscious creatures like ourselves, whereas the present Fundamental Singleness claim doesn't depend on the subject's having any sort of express or 'thetic' experience of itself as subject, let alone any such experience of itself as single considered specifically as subject. If worms have experience, their experience has Fundamental Singleness, at any given time, and not just by being very simple. So does the chaotic experience of schizophrenics, considered at any given moment.

There's another respect in which Fundamental Singleness is more basic than the claims about unity for which Kant is famous. What is in question is a necessary unity of experience or experiential field that we can know to exist before we've had any distinctively Kantian thoughts about the 'necessary unity of consciousness' or 'necessary synthesis of the manifold'.[42] It is, in other words, a necessary unity of experience that we can know to exist before we come to any aspect of the unity of experience that we can know to exist because it's a necessary condition of the possibility of other relatively complicated things that we know to exist – *e.g.* articulated thoughts like *grass is green*; or experiences that involve 'co-consciousness' in being at once and indissolubly both visual and olfactory and auditory; or what Kant calls 'experience', by which he means what some now call 'objective experience', *i.e.* experience that has, for the experiencer, the character of being experience of a world of objects existing independently of the experiencer. What is in question is a still more basic necessary unity of experience or experiential field we know to

[41] *Critique of Pure Reason*, translated by N. Kemp Smith (London: Macmillan, 1781–7/1933), A355. If one moves to a notion of apperception (*i.e.* an entity's awareness of its own mental being) that allows apperception to entities that lack anything like full or express self-consciousness, then the two claims may – perhaps – come together again; but even then there is a natural way of taking the notion of apperception given which the present point holds completely independently of any questions about apperception.

[42] '… the necessary unity of consciousness, and therefore also of the synthesis of the manifold' (*Critique of Pure Reason*, *op. cit.*, A109; this is the only occurrence in the *Critique* of the phrase 'necessary unity of consciousness').

exist not by running any kind of transcendental argument of this kind but simply by considering what it is for a subject to be having any experience at all – by considering what it is for there to be an experiential point of view at all.[43]

I think, then, that the Fundamental Singleness point is secure, and it gives me the strong unities I need. But I'm now going to call two of my heroes – Descartes and Kant – in further support.[44]

10. Descartes on Fundamental Singleness

In the case of Descartes, a single point. When Descartes claims that the indivisibility of the mind or soul or subject is something certain,[45] Fundamental Singleness is a key part of what he has in mind. He has been widely condemned (even mocked) for this indivisibility claim, but it's certainly correct given the central tenet of his metaphysics of mind. His conception of the mind or soul is nothing like the standard immaterialist conception of it as some sort of locus of consciousness or 'thinking' that isn't itself wholly constituted of consciousness or thinking. His truly radical view is that *res cogitans* – mind, soul – just is (necessarily subject-of-experience-involving) experience or consciousness or thinking; it has no other mode of being. This being so, the indivisibility thesis is nothing more than the assertion of Fundamental Singleness, at least so far as the synchronic case is concerned. (Kant doesn't criticize this view in the 'Paralogisms' section of the *Critique of Pure Reason*; he simply criticizes a certain metaphysical use to which it is put.)

11. Kant on Fundamental Singleness

In the 'Transcendental Deduction' and 'Paralogisms' sections of the *Critique of Pure Reason* Kant discusses experiential unities that are

[43] Note that I take it as given that things just are a certain way for one, experientially, at any given moment of experience, and that it is entirely determinate how they are, even if – even though – we cannot specify in exhaustive detail how they are. See G. Strawson, 'Can we Know the Nature of Reality as It is in Itself?' in *Real Materialism and Other Essays*, *op. cit.*, 77–8.

[44] In *Selves* I also recruit William James; see especially *The Principles of Psychology*, Volume 1 (New York: Dover, 1890/1950), 276–8; and 371, where he argues for the 'indecomposable unity of every pulse of thought'.

[45] In his 'Sixth Meditation', in *The Philosophical Writings of Descartes*, *op. cit.*, Volume 2, 59.

the thinking of thoughts in the narrower, cognitive, non-Cartesian sense, particular events of proposition-comprehension, *e.g.* judgements. We can, I propose, discern – or at least extract – the following unity argument.[46] We know, to begin, that

(1) particular thoughts exist.

We also know that

(2) such thoughts must have distinguishable elements, or at least a certain complexity, in order to be genuine thoughts at all, discursively articulated thoughts like *grass is green*.

We also know that

(3) these distinguishable elements, or this complexity, must form a unity, the (mental) unity that is the phenomenon of the comprehension of the proposition by a subject, if there is to be a genuine thought at all.

On the present materialist (um-ist) view

(4) such unities are wholly physical unities, in being mental unities, although they may not present as intuitively unitary entities when non-experientially considered (*e.g.* when examined by neurologists).[47]

We know, then, that there are actually existing physical entities that concretely realize a certain sort of unity. This (interim) conclusion is not very exciting, because the same can be said of cars or bananas.[48] But the further point is that the concrete unity in question is a unity of a very strong sort. It's the 'logical unity of every thought'.[49] Cars and bananas can lose car parts or banana parts and retain their identity as the cars or bananas they are, so long as we're reasonably relaxed about the identity conditions of things. A thought, by contrast, can't lose any thought-part and retain its identity as the particular thought it is, and the same goes for a total experiential field. Every part is as essential as every other part just in being, indeed, a part. They are indivisible (and in that sense simple) unities.

If we accept this, we can infer that

[46] The main argument occurs in Kant's Transcendental Analytic, but can be easily integrated with his discussion of the 'I' or self in the Paralogisms section of the Transcendental Dialectic, and I move freely between the two.

[47] I. Kant, *Critique of Pure Reason*, *op. cit.*, A355.

[48] As Darragh Byrne pointed out in discussion.

[49] I. Kant, *Critique of Pure Reason*, *op. cit.*, A398.

(5) there are actually existing physical entities – thoughts – that concretely realize a certain sort of very strong, 'logical' unity.

What is this unity? It's the unity specified in (3), nothing more; the unity that must exist if a thought is to occur at all. It is, on the one hand, a very ordinary everyday thing, but it is also, on the other hand, something completely exceptional, because it's an absolute, concretely existing unity.

There's a further reason why it's exceptional. If we ask generally what logical or absolute unity or indivisibility might be, a first thought is that it involves perfect simplicity. On this view, if something is an absolute unity, then

[U] it has no true description that furnishes any sense in which it can be said to involve complexity, or parts.

The proposed absolute or logical unity of a thought such as *grass is green* is not like this, however, for

[¬U] it has a true fundamental description that displays it to have a certain sort of complexity

(the same goes, presumably, for almost any total experiential field). If this is right, then even if [U] is sufficient for absolute unity or indivisibility, it isn't necessary. One might call absolute unities/indivisibilities which fail to satisfy [U] 'complex absolute unities', and one might then wonder whether there can be such things. I am at present supposing that there are many of them, however mysterious they may seem considered under this description, but one might also suppose that there's only one – the universe. This, though, raises the question of the one and the many, which I don't want to consider. (The best answer to the question, perhaps, given discursive thought's limited capacity to represent the nature of reality, is that it is true that there is only one thing and true that there are many things.)

The heart of Kant's notion of logical unity is in any case simple. It is in fact the Fundamental Singleness point, adjusted to the particular case of thought. If a thought is had, if a proposition is genuinely entertained, then it's necessarily had or entertained by a single subject. In Kant's words,

(6) the 'logical unity of every thought' is inseparable from the 'absolute … logical unity of the subject' of the thought.[50]

[50] *Ibid*., A356.

And it's not only true that concretely occurring thoughts realize a certain sort of unsurpassable concrete unity. It's equally true that

(7) the (thin) subjects of thoughts – concretely realize a certain sort of 'absolute … logical' unity.

As Kant observes:

> That … the *I* in every act of thought is *one*, and cannot be resolved into a plurality of subjects, … is something that lies already in the concept of thought, and is therefore an analytic proposition.[51]

I believe that these two unities are in fact and at bottom the same unity, and take it again that Kant agrees, in a passage already quoted, when he says that 'the thinking or the existence of the thought and the existence of my own self are one and the same'. But this is a further, difficult idea. For the moment we may take it that when we have to do with these actual concrete logical or absolute unities we have to do with unities of an absolute or unsurpassable sort, unities that must count as strong unities on any account of what strong unity is.

So in knowing that such thoughts really exist, as I take it we do, and in knowing that any thought must have a certain sort of absolute unity if it is to exist at all (even though there's also a respect in which it must involve difference and complexity), we know – to put it in terms that may seem provocative to some Kantians, but shouldn't seem so to Kant, so long as he grants that thoughts really concretely exist – that there exist entities that are genuine, *concrete*, *metaphysical* unities or indivisibilities of an unsurpassable sort. (Obviously I have to drop 'physical' and use 'concrete' in order to state the point in a way Kant can accept.)

The same degree of unity exists in the case of multimodally complex sensory experiences, and a parallel argument can be constructed for the necessary singleness of the subject of any such experience.[52] The same degree of unity also exists in the case of experiences that are absolutely simple in content, of course – if such experiences, sensory or otherwise, are possible.[53] The only difference is that in this

[51] *Ibid*., B407.

[52] Kant's intellectualism may suffice to explain any privileging of thoughts over sensory experiences, but it's also important that he explicitly allows for the possibility that there could be thought without sensory experience, *i.e*. 'intellectual intuition'.

[53] We might allow a sense in which a pure note is simple although it involves pitch, timbre and loudness, and there seems no reason in principle why a sensory state-space could not be strictly one-dimensional. 'Pure

case we lack the vivid illustration of the force of the unity that derives from the respect in which there is absolute unity in spite of complexity.[54]

Here, then, or so I propose, we move validly (*i.e.* without paralogism) from a 'logical' or 'logico-phenomenological' point to a substantive metaphysical conclusion. There may be neurological or cognitive-science accounts of what's going on that find astonishing complexity in these cases, many disparate and discrete events. None of this matters as long as there's a genuine, concrete *experiential* unity – *i.e.* the having of a thought, the actual occurrence of a thought. For, crucially, this unity is itself a concretely existing metaphysical unity or indivisibility of an unsurpassable sort, the unity of the concrete event of the genuine entertaining of a thought.

The phrase 'experiential unity' mustn't be misunderstood. There is as already remarked no requirement that the subject have any express experience of itself or its thought *as* a unity. The requirement is only that it have a genuine thought – that a genuine thought occur. This is the experiential unity in question, which is itself an unsurpassable, concrete metaphysical unity.

It may be granted that Kant is committed to the existence of thoughts and judgements, by his moral views if by nothing else, but objected that he can't and won't claim to know that these existent thoughts are metaphysically unsurpassable unities, if only on the grounds that they have an irreducibly temporal character, so far as we know anything about them, and so can't be known by us as they are in themselves, but only in so far as they appear to us. Against this we may bring Kant's claim that they – and their subjects – involve 'absolute' unity. I think that Kant would – must – on this ground be prepared to agree that they're unsurpassable unities that we can know to exist and to be such. My present aim, though, is to make a proposal, not to attempt detailed Kantian exegesis.

> – I really don't think you can expect Kant's blessing. What about the passage in which Kant speaks of the 'merely logical qualitative unity of self-consciousness in thought in general, which has to

---

awareness', of the sort that Buddhists claim to attain, is another possible example of absolute simplicity of content.

[54] There are cases in which it's natural to say that someone has had or has entertained *a* thought, but in which what we mean by 'a thought' is such that having or entertaining it involves more than one unity of the sort I have in mind at present.

> be present whether the subject be composite or not'?[55] His point here is precisely that there may be no real substantial metaphysical unity at all although there may be and indeed must be this 'qualitative unity'.[56]

Kant has a very specific *ad hominem* purpose in the Paralogisms. His aim is to stop the [simple/single → indivisible → immaterial → incorruptible → immortal] chain of metaphysical argument that was popular in his time.[57] In this traditional sense of 'metaphysical' Kant's claim is indeed, quite unequivocally, that we have no good grounds for believing that there's a real metaphysical singularity or simplicity or unity. This is his correct objection to the use that the Cartesian 'rational psychologists' aimed to make of Fundamental Singleness (it's not original to him). The present proposal, though, is that there's another wider sense of 'metaphysical' that Kant not only can but seemingly must allow, which yields a real, known metaphysical unity after all,[58] and doesn't disrupt his argument against the rational psychologists. In this way we legitimize the Second Paralogism argument by removing the equivocation on the word 'subject' that renders it paralogistic: the subject that is knowably single is also knowably a metaphysically real concretely existing entity.

The proposal depends as before simply on the claim that Kant agrees that thoughts – experiences, judgements – really, concretely exist (I avoid the explicitly temporal word 'occur'); that they are indeed metaphysically real. It goes from there to the point that he must then agree that the unity that their existence essentially involves must truly exist and be real, a genuine metaphysical unity of some

[55] I. Kant, *Critique of Pure Reason*, *op. cit.*, B413.

[56] Compare, also, the use of 'qualitative' in the Transcendental Deduction on B114.

[57] 'Everyone must admit that the assertion of the simple nature of the soul is *of value* only in so far as I can thereby distinguish this subject from all matter, and so can exempt it from the dissolution to which matter is always liable. This is indeed, strictly speaking, the only use for which the above proposition is intended' (*Critique of Pure Reason*, *op. cit.*, A356; my emphasis). There is a respect in which we may 'allow full objective validity' to the 'proposition … *everything which thinks is a simple substance*' (A357), but 'we still cannot make the least use of this proposition in regard to the question of [the mind or soul's] dissimilarity from or relation to matter' in such a way as to hope to establish immateriality and thence incorruptibility and immortality.

[58] It is known that it exists, and that it is a unity, whatever the sense in which its nature remains unknown.

sort, although it licenses absolutely no conclusions about immortality and so forth.

It may be objected that the unities in question (unity of subject, unity of thought or experience) are 'merely functional' unities, but it's not clear what this 'merely' means, as applied to concretely actualized entities (a concretely existing total experiential field, and a concretely existing subject for whom that field is an experiential field). In particular, it's not clear that it detracts in any way from these unities' claim to be true metaphysical unities. Even if we can find a sense in which they're 'merely' functional unities, this won't touch the fact that they're concretely existing unities, genuine metaphysical unities. And by the time we've revised and processualized the notions of object and substance in the way we have to if they're to retain any place in fundamental metaphysics, and stripped them of their claim to stand in fundamental ontological contrast with attributes or properties,[59] we'll be in no position to hold on to any view that has the consequence that these unsurpassable thought unities or experience unities, these thinker unities and experiencer unities, have less claim to be objects or substances than paradigmatic examples of ultimates (or indeed good old fashioned immaterial minds).

We may as materialists suppose that the existence of these unsurpassable mental unities involves the existence of a large number of neurons (and *a fortiori* ultimates) acting in concert. In this case we may say that it essentially involves a plurality of 'different substances acting together', to use Kant's general formulation in his discussion of the Second Paralogism.[60] Kant is right that we can't ground any claims about the non-materiality and hence possible simplicity or indivisibility and hence incorruptibility and hence immortality of the soul on the knowable existence of these 'absolute' concrete unities of subject and experiential field.[61] This, again, is the line of thought he demolishes in his discussion of the Paralogisms. This demolition is, however, wholly compatible with, and, crucially, *derives its principal force from*, the fact that we can know that these fundamental and unsurpassable unities exist. We can know they exist even though they're of no use in establishing the existence of a persisting immaterial subject. It's Kant himself, as remarked, who

[59] I argue for this in Strawson, *Selves*, *op. cit.*, 299–320.

[60] I. Kant, *Critique of Pure Reason*, *op. cit.*, A353.

[61] Kant does, however, take his arguments in the Paralogisms to show that materialists can't give an adequate explanation of the existence of the subject, given its 'logical' simplicity property. See *ibid.*, B419–20.

makes it most vivid that the fact that 'the I in every act of thought is *one* … lies already in the very concept of thought'. He's prepared to '*allow full objective validity*' to the '*proposition … everything which thinks is a simple substance*', at least for the sake of argument, for his point is that even if we do this 'we still cannot make *the least use* of this proposition in regard to the question of [the mind or soul's] dissimilarity from or relation to matter' in such a way as to hope to establish immateriality and thence incorruptibility and immortality.[62] We can know the necessary singleness or simplicity of the subject of experience in the having of experience, but to know this is not to have 'knowledge of the simple nature of the self as subject, such as might enable us to distinguish it from matter, as from a composite being'.[63]

So much for the 'merely functional unity' objection. It may now be objected that Fundamental Singleness is a 'merely experiential unity' and so not a real metaphysical unity. In one version, the objection is that the claim that the total experiential field of an experience is itself something that is a true metaphysical unity is unwarranted, because this unity is really only an appearance – because it's only from a certain point of view that the total experiential field presents as unified. In a second version the objection is aimed directly at the (thin) subject of the experience, the objection being that the unsurpassable unity of thin subject, too, is, somehow, 'merely experiential' or 'subjective', and so not ultimately real.

There's a lot to say to this, but the basic reply is simple: what's 'merely' about this unity? It's true that the unity of the total experiential field of an experience is essentially, and if you like 'merely', an *internal experiential* or *IE* unity, whatever else it is or isn't. It's true that it's a unity considered just in respect of its internal experiential character, a unity considered from the point of view of its subject, a unity that is in fact constituted as such by the existence of the point of view of its subject.[64] It's true that there is a crucial sense in which its fundamental-unity characteristic is wholly a matter of its experiential being (when we consider the existence of the experience non-experientially, we find millions of ultimates in a certain complex state of interaction). All this is true. But it's no less true that it's an objective, ground-floor metaphysical fact that subjective experience exists and that it has the character it does. That is: the experiential

[62] *Ibid*., A357; my emphasis.
[63] *Ibid*., A360.
[64] However unreflective the subject, however little 'thetic' apprehension it has of the fact that the unity in question is a unity.

fact that one of these IE unities exists is itself an objective ground-floor metaphysical fact, a fact about the world. It's a fact that stands as a fact from a perspective external to the perspective internal to the IE unity. The reason why we're right to judge an IE unity to be a fully metaphysically real unity when considering it *externally*, from the outside, objective point of view, is indeed nothing other than the fact that it is indeed a strong *internal* experiential unity, but there's nothing odd about the sense in which subjective phenomena can be objective facts. This is part of any minimally sensible realism about the subjective. All that's added here is the slightly more specific point that the existence of an IE unity that consists of the experiencing of a total experiential field by an experiencing subject at a given moment is itself an objectively existing real unity (experience is real!), both in its single-subject aspect and in its single total-experiential-field aspect (which may at bottom be the same thing). The statement that this unity exists and is wholly metaphysically real holds true outside the perspective of the subject in question, although the unity in question consists wholly of the phenomenon of the existence of that subject having that experience, with the content that it has, at the time in question.[65]

12. Conclusion

One can dispute all sorts of details, but the Fundamental Singleness point and Kantian points remain as secure. The unity of the thin subject and the experiential unity of its experience are two aspects of the same unity. There is, as remarked at the beginning, no explanatory relation between them. The unity in question is by hypothesis a concrete unity, and I've argued that it is an exemplary strong unity, an unsurpassable example of a unity:

[20] [subject → strong unity].

[65] The qualification one then needs to enter is that there need be no distinctive *experience of unity* on the part of the subject of the IE unity, inasmuch as the necessary unity in question is simply the fact already mentioned: the trivial fact that any subject necessarily experiences all the material it experiences at any given time in a single experiential perspective. Some Phenomenologists may hold that there must in every case, however lowly, be some sort of experience of unity *experienced as such*. I think I can agree to this, in any sense in which it is true, without disturbing the point of this qualification.

Given, then, that strong unity is the fundamental – and sufficient – criterion of objecthood, *i.e.*

[18] [strong unity → object],

it follows that

[21] [subject → object]

– that concretely existing subjects of experience, which we know to exist, are objects in the fullest metaphysical sense. More moderately, I've argued that there are no better candidates in nature for the title 'strong unity' – at least until we give up the struggle to maintain that there is a plurality of objects, and revert to the lovely Spinozan position that there is only one, 'Nature', *i.e.* the Universe. If there are any objects at all, then subjects are objects. In fact they're the best examples of strong unities that we have. Any metaphysics that has trouble with this point needs an overhaul.

University of Reading

It Must be True – But How Can it Be? Some Remarks on Panpsychism and Mental Composition

PIERFRANCESCO BASILE

Abstract
Although panpsychism has had a very long history, one that goes back to the very origin of western philosophy, its force has only recently been appreciated by analytic philosophers of mind. And even if many still reject the theory as utterly absurd, others have argued that it is the only genuine form of physicalism. This paper examines the case for panpsychism and argues that there are at least good *prima facie* reasons for taking it seriously. In a second step, the paper discusses the main difficulty the theory has to face, the 'composition problem'. This is the problem of explaining how the primitive experiences that are supposed to exist at the ultimate level of reality could give rise to the unified experience of a human being. What assumptions as to the nature of experience generate the composition problem? Is mental composition impossible in principle or do we simply lack at present any understanding of phenomenal parts and wholes?

1. Introduction

Analytical philosophy of mind has recently witnessed a revival of interest in the old-fashioned doctrine of panpsychism – the view that experience is a pervasive feature of reality. The doctrine is not uncontroversial; while some embrace it almost enthusiastically, others reject it as outrageously absurd. On the one hand, there are supporters like David Skrbina, who argues that 'panpsychism… *offers resolutions* to mind-body problems that dualism and materialism find intractable'.[1] On the other end of the scale, one finds pitiless critics like Colin McGinn, who compares panpsychism unfavourably with Cartesian dualism and belief in divine providence, arguing that the theory is not even a form of supernaturalism – it is 'merely extravagant'.[2]

As it will be argued, the right attitude should be neither one of condemnation nor one of uncritical endorsement. There are wholly

[1] D. Skrbina, *Panpsychism in the West* (Cambridge, Mass.: The MIT Press, 2005), 4; my emphasis.

[2] C. McGinn, *The Problem of Consciousness: Essays Towards a Resolution* (Oxford: Blackwell, 1991), 2n.

doi:10.1017/S1358246110000044

cogent lines of reasoning conducive to it; at the same time, its viability is threatened by a few as yet unanswered questions. To put it in a way more favourable to the panpsychist: the doctrine cannot be taken to provide a fully satisfactory explanation of the place of consciousness in nature; it does have its own strengths, however, and does therefore deserve a fair hearing.

2. Physics, Metaphysics, Commonsense

John Searle provides a brief criticism of panpsychism in his book of 2004, *Mind: A Brief Introduction*. His assessment is as good a point of departure for an evaluation as one could possibly wish. The present paper can be read as a response to what he says there:

> Panpsychism is the view that consciousness is everywhere. This view is seldom stated explicitly, but it is in several authors particularly among the mysterians who think that if we are going to explain consciousness in terms of microprocesses, then, somehow or other, some form of consciousness must already be present in the microprocesses. … On this view, everything is conscious to some degree. In giving an example of the ubiquity of consciousness, Chalmers eloquently describes what it might be like to be a conscious thermostat.[3]

After this brief – and hardly sympathetic – characterization of the theory, Searle launches his attack:

> Aside from its inherent implausibility, panpsychism has the additional demerit of being incoherent. I do not see any way that it can cope with the problem of the unity of consciousness. … If the thermostat is conscious, how about the parts of the thermostat? Is there a separate consciousness to each screw? Each molecule? If so, how does their consciousness relate to the consciousness of the whole thermostat? And if not, what is the principle that makes the thermostat the unit of consciousness and not the parts of the thermostat or the whole heating system of which the thermostat is a part or the building in which the heating system exists?[4]

[3] J. Searle, *Mind: A Brief Introduction* (Oxford-New York: Oxford University Press, 2004), 149–50.

[4] *Ibid*., 150.

This passage nicely summarizes what can be said – and what has in fact been said – against panpsychism; one will have to disentangle, however, Searle's misunderstandings from his genuine insights. Let's begin with what is most obviously unsatisfactory, namely his rejection of panpsychism on the ground of its 'inherent implausibility'.

Charges of implausibility are notoriously hard to evaluate – by what standards is the plausibility of a philosophical theory to be tested? Searle does not say. The standard can hardly be common-sense; if it were, this would suffice to reject all the –isms in the philosophy of mind (perhaps, only such a vituperated doctrine as Cartesian dualism would avoid such a wholesale condemnation!). Be that as it may, it is difficult to see why panpsychism should be regarded as more absurd than any of the existing forms of materialism, which enjoy on the contrary a very wide acceptance.

Is it 'science', then, that provides the required standard? The problem here is that it is unclear what this is supposed to mean. Should one include, say, physics, chemistry, biology, or only one of these? And why exactly is science supposed to rule out panpsychism? Appeals to 'our best scientific theories' are common in the philosophy of mind, yet quite often they are little more than empty slogans: I suspect that there is more than a grain of truth in Ross and Ladyman's critique of recent works in analytic philosophy of mind as being based upon a 'pseudo-scientific metaphysics'.[5]

Most importantly, panpsychists argue that natural science does not make any positive claim about the *intrinsic* nature of reality; rather, they claim that science limits itself to the study of the *formal* properties of the investigated objects. As Russell says with regard to our knowledge of matter:

> [I]t is only mathematical properties that we can discover… The physical world is only known as regards certain abstract features of its space-time structure – features which, because of their abstractness, do not suffice to show whether the physical world is, or is not, different in its intrinsic character from the world of mind.[6]

[5] D. Ross and J. Ladyman (with D. Spurrett and J. Collier), *Every Thing Must Go: Metaphysics Naturalized* (Oxford: Oxford University Press, 2007), 18–27.

[6] B. Russell, *Human Knowledge: Its Scope and Limits* (London: Routledge, 1948), 240.

Thus, so the argument goes, there cannot possibly be anything in science that contradicts panpsychism: 'physics' and 'meta-physics' operate at *distinct* explanatory levels. True, the notion that science has nothing to say about the ultimate nature of things is not an obvious one, but the argument needs to be addressed in its own terms – a generic appeal to implausibility will not suffice.

Another noteworthy feature of the passage is Searle's contention that panpsychism is 'seldom stated explicitly'. Doesn't this suggest that panpsychists are a little dishonest, afraid, as it were, to clearly come into the fore? This can't be right; the list of philosophers *explicitly* committed to some form of panpsychism is not as brief as Searle would make us believe. The doctrine has an impressive history, one which includes names as venerable as those of Gottfried Wilhelm Leibniz, Benedict de Spinoza, Charles Sanders Peirce, William James, and Alfred North Whitehead. In recent times, the likes of philosophers such as David Chalmers, William Seager and Thomas Nagel have argued that it is a hypothesis worth contemplating. And it has been forcefully advocated by Charles Hartshorne, David Ray Griffin, David Skrbina, Galen Strawson, and Timothy Sprigge.

Ipse dixit? This is not, of course, an appeal to authority; nor is it simply a pedantic attempt at setting the record straight. The point is that panpsychism is *not* to be viewed as an historical anomaly; it should not, therefore, be dismissed as a philosophical eccentricity.

3. The Ubiquity of Experience

This might not be enough to dispel the suspicion that the theory is implausible – if not even obviously absurd. Searle touches upon a serious problem at the very end of the quoted passage. What he has in mind is the specific version of panpsychism that has been tentatively advocated by David Chalmers in *The Conscious Mind*. On this version of the theory, a thermostat has a claim to be regarded as sentient. But a thermostat is made up of parts (the screws, the molecules), and is itself part of larger complexes (the whole heating system, the building). Quite rightly, Searle asks why only the thermostat should be regarded as sentient. Where – and, most importantly, *how* – do we draw the line between the sentient and the non-sentient?

In the case of Chalmers' version of panpsychism, the answer is straightforward. Since he speculates that experience might be linked with information, whatever can be regarded as an information

system could also be regarded as endowed with experiences.[7] The real issue here is to account for the nature of the link that connects information and experience; nothing is truly explained by arguing that experience and information are like the two sides of a manuscript – *recto* and *verso*.

In general, two lines of response are available to the panpsychist. (1) One is to ascribe experiences to everything that exists. In a recent essay, David Skrbina argues that there is no difference in kind between a human being and a rock ('*the rock experiences the world*, and *we experience the world*'), before going on to ascribe a mental life to 'all things that exist – from atoms and rocks, to tables and chairs, to human beings, planets, and stars'. This is legitimate, but it isn't very satisfying; it simply doesn't ring true that the theory 'finds justification in… empiricism'.[8] There is nothing in our experience that even slightly suggests that sticks and stones, tables and chairs, planets and galaxies, are experiencing things. As John Dewey recognized, '[t]here is no evidence that experience occurs everywhere and everywhen'.[9] On the contrary, at least in some cases, our experience suggests (although it does not prove) the opposite view. Stones are incapable of spontaneous movement and reaction, which clearly demarcates them from those organisms – such as human and non-human animals – to which we do not hesitate to ascribe a conscious, experiential life. Although empirical observation does not show that sentience is not ubiquitous, we should treat its deliverances as we are supposed to treat a literary text – that is to say, *we should not read too much into them*.

This critique differs radically from Wittgenstein's remark in the *Philosophical Investigations*: 'Could one imagine a stone's having consciousness? And if anyone can do so – why should that not merely prove that such image-mongery is of no interest to us?'[10] The point is not that a panpsychism that goes so far as to ascribe mental states to such things as sticks and stones would be meaningless; we can make *some sense* of such a proposal. Simply, such a version of

[7] D. Chalmers, *The Conscious Mind* (Oxford-New York: Oxford University Press, 1996), 293.

[8] D. Skrbina, 'Transcending Consciousness: Thoughts on a Universal Conception of Mind', *Journal of Consciousness Studies* **16:5** (2009), 81 and 84.

[9] J. Dewey, *Experience and Nature* (New York: Dover, 1958), 3a.

[10] L. Wittgenstein, *Philosophical Investigations*, edited by G. E. M. Anscombe and R. Rhees (Oxford: Blackwell, 1953), §390.

panpsychism would assume more than is empirically warranted[11] and, arguably, more than is needed for explanatory purposes.

(2) This becomes clearer if one considers the other option which remains open to the panpsychist. Specifically, it is possible to differentiate between wholes which, considered as such, are capable of sentience from wholes which, considered as such, are deprived of experience. The panpsychist might then choose to qualify the notion that experience is ubiquitous by making the following proposal (with 'E' and 'NE' standing for 'experiential' and 'non-experiential' respectively):

> [A] All ultimate constituents of reality are sentient.
> [B] When such ultimates are grouped in certain special ways – say, the *E*-ways – they give rise to complexes that are themselves sentient.
> [C] When they are grouped in different ways – say, the *NE*-ways – they give rise to complexes that are themselves insentient.

This would not be implausible. There is no reason to believe that a complex must be sentient solely because its parts are; this would be an instance of the *fallacy of composition*. Conversely, there is no reason to believe that there is no experience in the parts, solely because the whole is not itself an experiencing thing; this would be an instance of the *fallacy of division*. There is nothing that suggests permanence and immovability more vividly than a solid piece of rock, yet we have no difficulty in accepting the scientific description of it as composed of vibrating, energizing particles. It is at least conceivable that the same could be true of the sentient/insentient distinction.

In order to be entitled to hold a position of this sort, however, the panpsychist must be able to explain what 'principle' accounts for the fact that only certain wholes posses a unified consciousness. One would need a clear account of the distinction between those ways of organization that give rise to sentient wholes (the *E*-ways) and

[11] One truly major thinker who went so far as to ascribe experiences to all things is Spinoza. It is significant that this claim was not established empirically, but was inferred as a corollary of a general metaphysical theory established on *a priori* grounds. Specifically, given his substance metaphysics and the parallelism that goes with it, every mode of extension has its counterpart in a mode of thought. Thus, there must be of necessity a mental correlate to each physical thing. His metaphysical principles, he argues, 'are completely general and do not pertain more to man than to other individuals, all of which, though in different degrees, are nevertheless animate'. B. Spinoza, *Ethics*, edited and translated by E. Curley (London: Penguin Books, 1994), 2p13.

those that do not (the *NE*-ways). Such a theory would be less difficult to digest, but it would be a hard task to work it out. A speculative philosopher such as Whitehead appears at times to be proposing a tremendously intricate taxonomy of complex individuals, although only in a sketchy manner.[12] To the best of my knowledge, the detailed articulation of such a theory has yet to be provided.

4. The Heterogeneity Problem

The above considerations show that the panpsychist need not be committed to the view that apparently inert objects such as rocks or thermostats have experiences. In order to dispel fully the suspicion that the theory is absurd, however, one has to consider the reasons in its support. Searle condemns the view as a form of 'mysterianism' but doesn't make the slightest effort to explain why very intelligent philosophers should have come to entertain such an apparently strange idea in the first place.

Although a variety of arguments has been advanced over the centuries, the one that is likely to get a hearing in the current philosophical climate pivots upon the notion of emergence. James provides a vivid illustration of the perplexity that grounds the panpsychist's case in his *Principles of Psychology*:

> The point which as evolutionists we are bound to hold fast to is that all the new forms of being that make their appearance are really nothing more than results of the redistribution of the original and unchanging materials. The self-same atoms which, chaotically dispersed, made the nebula, now, jammed and temporarily caught in peculiar positions, form our brains; and the 'evolution' of the brains, if understood, would be simply the account of how the atoms came to be so caught and jammed. In this story no new *natures*, no factors not present at the beginning, are introduced at any later stage.
>
> But with the dawn of consciousness an entirely new nature seems to slip in, something whereof the potency was *not* given in the mere outward atoms of the original chaos.[13]

[12] A. N. Whitehead, *Process and Reality: An Essay in Cosmology*, corrected edition by D. R. Griffin and D. W. Sherburne (New York: The Free Press, 1978), 83–109.

[13] W. James, *The Principles of Psychology*, Vol. I. (New York: Dover, 1950), 146.

The obvious conclusion is that in an evolutionary universe experience must be present from the very beginning. One of Russell's teachers at Cambridge, James Ward, argued as follows:

> It is interesting… to notice that in the support which it lends to pampsychist views the theory of evolution seems likely to have an effect on science the precise opposite of that which it exercised at first. That was a leveling down, this will be a leveling up. At first it appeared as if man were only to be linked with the ape, now it would seem that the atom, if reality at all, may be linked with man.[14]

This is a truly remarkable passage and its power should not pass unnoticed. Even today – and not solely in the popular press – evolution is associated with a materialistic conception of reality; in orthodox circles, it is even rejected on this very ground. If Ward is right, however, then the link between evolution and materialism might not be as close as we are accustomed to think. The theory of evolution need not necessarily degrade consciousness, for *it might spiritualize matter*.

In remarking that the generation of the experiential from the non-experiential would involve a break of evolutionary continuity, James and Ward are raising the 'heterogeneity problem'. This is a difficulty Galen Strawson has recently stated in very clear terms.[15] As he puts it, the problem arises for a philosopher inclined to hold the following two theses (with 'NE' and 'RE' standing for 'non-experiential' and 'real physicalism' respectively):

> [NE] physical stuff is, in itself, in its fundamental nature, something wholly and utterly non-experiential.
> [RP] experience is a real concrete phenomenon *and* every real concrete phenomenon is physical.

The problem here is to reconcile the belief that mental phenomena are physical, [RP], with our deep ingrained intuition that physical stuff

[14] J. Ward, 'Mechanism and Morals: The World of Science and the World of History' in *Essays in Philosophy: with a Memoir by Olwen Ward Campell*, edited by W. R. Sorley and G. F. Stout (Cambridge: Cambridge University Press, 1927), 247.

[15] G. Strawson, 'Real Materialism: Why Physicalism Entails Panpsychism' in *Consciousness and its Place in Nature: Does Physicalism Entail Panpsychism?*, edited by A. Freeman (Exeter: Imprint Academic, 2006), 12–21.

isn't experiential, [NE]. Most philosophers would surely wish to believe both. Yet it is not easy to see how they could be entitled to do so, for if [RP] is true, then some real things are both physical and experiential, which is what [NE] denies.

The way out of this *impasse* is supposed to be provided by the notion of emergence. When purely physical stuff achieves a certain degree of complexity in its organization, then new properties – experiential properties – come into existence. In this way, it becomes possible to hold that all physical stuff is, in its fundamental nature, non-experiential [NE], without having to deny that experiential phenomena are wholly physical [RP], since they are now conceived as the outcome of wholly natural processes. In a book entitled *Mind and Emergence*, for example, Philip Clayton characterizes emergence as 'the view that new and unpredictable phenomena are *naturally* produced by interactions in nature'.[16] But is emergence an intelligible notion? On the face of it, the 'experiential' and the 'non-experiential' are utterly different categories of being: how could the purely physical, in the sense of the non-experiential, bring the experiential into existence? This does look like a miraculous feat – a kind of unfathomable *creatio ex nihilo*. As John Locke forcefully makes the point (although in a different argumentative setting): 'it is as impossible that incogitative matter should produce a cogitative being as that nothing, or the negation of all being, should produce a positive being or matter'.[17]

This is the whole point on which the panpsychist rests his case. If the notion of brute emergence is rejected, then one will be brought back to the original contradiction and will have to reconsider either [RP] or [NE]. One way to abandon [RP] would be to deny the reality of mental phenomena and embrace eliminative materialism. But the very existence of experiences does seem difficult to deny – what is more real than the pains I feel when hitting my foot against the table? Nor can it be that they simply *appear* to be, but are *really* not: since merely to appear is already to be a part of reality. (If there is a theory that has a fair claim to be regarded as an historical anomaly, isn't this eliminative materialism?)

Alternatively, one might hold that mental phenomena are real while at the same time reject the notion that they are physical. But this would yield a form of supernaturalism and a bifurcation of

[16] P. Clayton, *Mind and Emergence: From Quantum to Consciousness* (Oxford: Oxford University Press, 2004), vi; my emphasis.

[17] *An Essay Concerning Human Understanding*, abridged and edited by John W. Yolton (London: Dent, 1994), IV. X. 11.

reality into a material and a mental world. This is unsatisfactory, because the mental is just as natural as the material. Moreover, this solution reopens the vexed question: how could these two worlds intermingle and constitute the single unified universe in which we live? (*If* there is anything uncontroversial in the philosophy of mind, this is the notion that an understanding of mind's place in nature requires a *monistic* ontology, that is, one that recognizes only a single *kind* of basic stuff.)

It is tempting, at this point – the temptation *is* a strong one, and should be appreciated as such – to succumb to the siren song and to take the radical step of dropping [NE] out of board:

> [¬NE] physical stuff is, in itself, in its fundamental nature, NOT something wholly and utterly non-experiential.

This is tantamount to accepting some form of panpsychism. Isn't there any way to resist this conclusion? I admit that the notion of radical emergence is a difficult one. Strawson observes that, if brute emergence were possible, then we couldn't rule out other kinds of miraculous transitions such as the emergence of concrete phenomena from abstract ones.[18] I do not know how to prove that miracles do not occur; to admit that they do, however, is to give up on the attempt to *understand* consciousness' place in nature.

At this juncture, the following objection is likely to arise: 'Sooner or later, science will explain the emergence of the experiential out of the non-experiential: why, then, not suspend judgment and see what science will bring?'[19] This would be to miss the force of the panpsychist's case. We do not need to wait for science to dispel the riddle. The panpsychist does not claim that we do not know *at present* how to explain brute emergence; rather, the claim is that we already know that such a radical form of emergence is impossible – or, alternatively, that if brute emergence is real, then it is not something we can possibly understand.

Note that this is not to deny that science might one day explain consciousness' place in nature. If it will, however, this will be at the price of abandoning the notion of insentient matter and the conception of emergence that goes with it. To see this is to recognize that the difficulty is not empirical, but *conceptual* – the gap between the

[18] *Ibid*., 19; for a similar point, see T. Nagel, 'Panpsychism' in *Mortal Questions* (Cambridge: Cambridge University Press, 1979), 194.

[19] See P. Simons, 'The Seeds of Experience' in A. Freeman (ed.), *Consciousness and its Place in Nature*, *op. cit*., 148–49, for an objection along these lines.

experiential and the non-experiential is a *logical* one: mere acquisition of novel empirical data would not help unless one reframes the concepts that are used to interpret those data. Otherwise put, *it is our current concept of matter as that which is wholly and utterly non-experiential which makes it impossible to understand how mind, the experiential, could emerge from it*. Surely, the request that we revise our ordinary conception of matter is a radical one, but it is *much less* drastic than eliminative materialism or Colin McGinn's skeptical suggestion that there is something about our cognitive endowments that makes it impossible in principle to solve the mind-body problem.[20]

Thus, what seems worthy of consideration at this point is whether there is a way of avoiding panpsychism *while still rejecting the notion of a brute emergence of the experiential from the non-experiential*. Neutral monism might seem to offer a way of escape. Why not argue that reality's basic stuff is neither mental nor physical, and yet it is the ground out of which both experiential and non-experiential properties originate? On closer examination, one easily sees why this solution wouldn't do. If one rejects emergence on the ground of the heterogeneity problem, then one is inevitably committed to saying that mental properties cannot originate from a wholly non-mental source. The neutral substratum will have to possess, if not full-blown mental properties, at least some proto-experiential features; again, this is tantamount to admitting that some form of panpsychism must be true – that the seeds of experience lie deep in the ultimate nature of things.

Nevertheless, the attempt to avoid the panpsychist option by adopting neutral monism is not wholly misguided. The intuition that animates that proposal is that, at an ultimate level of explanation, reality might not be adequately described in terms of the twin notions of 'experiential' ('mental') and 'non-experiential' ('physical'). Perhaps one does not really have to choose between [NE] and [RP]? They could *both* be false in that they are an inadequate way to set up the stage in the first place.

This seems to be the solution recently favored by Colin McGinn: 'The only way out of this' – he says – 'is to hold out the hope of a third level of description… Brains have properties beyond those of

[20] As he has it, 'we are cut off by our very cognitive constitution from achieving a conception of that natural property of the brain (or of consciousness) that accounts for the psychophysical link…the felt mystery comes from our own cognitive limitations, not from any objective eeriness in the world'. C. McGinn, 'Can we solve the mind-body problem?', *Mind* **98** (1989), 359.

experience and those of basic physics and biology. These properties might then mediate between the other two sorts of property, offering some sort of unification of all three levels'.[21] On this view, one would have to redesign the conceptual landscape in which the whole debate takes place. Significantly, McGinn talks of a mere 'hope' that there might be such a 'third level of description'; the existence of such a level is nothing more than a logical possibility. Thomas Nagel goes a step further along this path, as he speculates that we might find the problem of consciousness intractable simply because we lack the adequate conceptual categories: 'We need entirely new intellectual tools, and it is precisely by reflection on what appears impossible – like the generation of mind out of the recombination of matter – that we will be forced to create such tools'.[22]

One would not hold to such a hope or aim at such grand conceptual revision, however, unless one had strong reasons for rejecting panpsychism. One main difficulty of the theory will be considered presently. For the time being, it should be emphasized that the argument based upon the heterogeneity problem *is* a powerful one. 'I suspect', Timothy Sprigge wrote, 'that our reason for believing that something is inherently impossible must always lie in the fact that we find it more and more difficult to conceive of it the more we understand the nature of the elements we are instructed to combine…That is, the nearer we approach intuitive fulfillment of the meaning of the expression said to express it the more we feel baffled'.[23] On the face of it, brute emergence looks like magic – isn't it so? The burden of proof lies on the emergentist here.

5. Mental Composition

Searle raises a serious question when he says that panpsychism can't 'cope with the problem of the unity of consciousness'. On the panpsychist view, he asks, how do the parts of a sentient whole 'relate to the consciousness of the whole'? For the sake of illustration, let's assume that the human brain (instead of Chalmers' thermostat) is composed of parts that are either sentient or wholly experiential in

[21] C. McGinn, 'Hard Questions. Comments on Galen Strawson' in A. Freeman (ed.), *Consciousness and its Place in Nature*, *op. cit.*, 98.

[22] T. Nagel, *The View from Nowhere* (Oxford: Oxford University Press, 1986), 52.

[23] T. L. S. Sprigge, *The Vindication of Absolute Idealism* (Edinburgh: Edinburgh University Press, 1983), 128.

nature. Searle's question now becomes: how do the lesser consciousnesses entering in the constitution of the brain coalesce so as to form the unified consciousness of a human being?

This well-known difficulty was raised by James in his *Principles of Psychology*. In that book, he compared the experiences in the neurons with the thoughts of several persons in a row:

> Take a sentence of a dozen words, and take twelve men and tell to each one word. Then stand the men in a row or jam them in a bunch, and let each think of his word as intently as he will; nowhere there will be a consciousness of the whole sentence.[24]

More recently, Philip Goff has made the example of a big pain composed of several little pains, as it were, by way of addition:

> Consider a physical ultimate that feels slightly pained, call it LITTLE PAIN 1. Consider then such slightly pained ultimates, LITTLE PAIN 1, LITTLE PAIN 2, *etc*., coming together to constitute a severely pained macroscopic thing, call it BIG PAIN. The pained-ness of each of the ultimates comes together to constitute the pained-ness of BIG PAIN: an entity that feels ten times the pain of each LITTLE PAIN. The severe pained-ness of BIG PAIN is wholly constituted by the slight pained-ness of all the LITTLE PAINS.[25]

Clearly, a collection of thoughts about single words does not amount to a *single* thought about a whole sentence; by the same token, an assemblage of little pains is not the same as a big pain, whose qualitative feel differs from that of any of its constituents or of all of them collectively. This is uncontroversial; equally uncontroversial, the transition from the experiences in the neurons to the human experience cannot be accepted as a brute contingency, but stands itself in need of explanation. This is no trivial point. Should this transition prove to be inexplicable, one will have to conclude that panpsychism is in no better position than any of the alternatives examined in the previous section, as it would face *its own version of the problem of emergence*.

But do we *know* that such a transition – from the low-level experiences to the higher human experience – cannot possibly take place? On the basis of what has been said so far, there is no reason to construe the composition problem as *a reductio ad absurdum* of the doctrine of

[24] W. James, *The Principles of Psychology*, *op. cit*., 160.

[25] P. Goff, 'Experiences Don't Sum' in A. Freeman (ed.), *Consciousness and its Place in Nature*, *op. cit*., 57.

mental combination and, by implication, of panpsychism. Specifically, the panpsychist need not be committed to the view that the lesser selves should bring about a larger consciousness in the same way in which, say, mosaic pieces come together to form the larger picture. Their mode of combination must be of an altogether different type. Surely, there is nothing in either Searle's or Goff's passages to warrant their contention that panpsychism is incoherent.

Quite recently, Goff himself has recognized that, *as stated above*, the criticism does not suffice to show that mental composition is impossible in principle: 'Contrary to views I have expressed in earlier work' – he now writes – 'I believe that the panpsychist can make good sense of subjects of experience summing, and hence can get round the combination problem'.[26] As against this, I *do* think that the notion of mental composition poses a very hard challenge for panpsychism. To see why this is so, however, an attempt must be made to formulate the problem in more precise terms.

6. Why is There a Problem?

What assumptions about the nature of experience generate the problem? Strangely enough, this question is rarely asked; it just seems to be taken for granted that we have an adequate grasp of the difficulty at issue. But consider the two passages quoted above. James' question is about *intentionality* – how could mental events, each of which has a different intentional object (a single word), generate a mental event with a new intentional object (a whole sentence)? Goff's is about *qualia* – how could mental events, each of which has its own 'what-is-it-like-feeling', generate a mental event that has its own distinct phenomenology (the feel of a big pain as opposed to that of a little one)? Without wishing to adjudicate the question as to the nature of the relationship between the two alleged 'marks' of the mental – 'intentionality' and 'subjectivity' – , it should be admitted that it is not immediately clear that these questions are the very same one or that the solution to the one will entail the solution to the other.

Let's here consider the problem in terms of *qualia*. After all, it is the subjective, qualitative dimension of conscious experience that

[26] P. Goff, 'Can the panpsychist get around the combination problem?' in D. Skrbina (ed.), *Mind that Abides: Panpsychism in the new millennium* (Amsterdam-Philadelphia: John Benjamins Publishing Company, 2009), 133.

has led contemporary philosophers to reconsider the heterogeneity problem, thus providing the motivation for panpsychism in the first place.[27] Once more, it is to James that we will turn for advice. In his later work, *A Pluralistic Universe*, he provides a most interesting evaluation of panpsychism. This is developed in the context of a discussion of the philosophies of Gustav Fechner and Josiah Royce. As James himself points out, however, what he says possesses a broader philosophical significance and a general critique of panpsychism can be extrapolated.[28]

Drawing freely from James' exposition, the problem might be construed as depending upon a few assumptions as to the nature of our experience:

> PHENOMENAL ESSENTIALISM – this is the view that, for an experience, to *be* is to *feel* a certain way. In the case of a pain, it seems pointless to draw a distinction between what the pain is *in itself* and the *way it feels*; the former – what the pain is 'in itself' – is wholly exhausted by the latter – its qualitative, felt dimension. As James forcefully puts it, in the case of experience, 'appearance' and 'reality' are one and the same.[29]

Besides the idea that an experience is nothing over and above the manner of its appearance – the idea that 'a mental fact' is 'just what it appears to be'[30] – there is another assumption that needs to be taken into account. This is one James would seem to endorse in the *Principles of Psychology*:

> PHENOMENAL HOLISM – this is the view that, within a person's total psychical whole, the nature of a single identifiable experience ('single', in the sense that it can be counted as 'one') is essentially determined by the other experiences occurring alongside it – synchronically – within the whole.

This principle involves a rejection of the traditional atomistic view of the mind, one typically exemplified by Hume's notion of a perception as a self-subsistent, substance-like entity that could, as a matter of

[27] Another limitation, already implicit in what has been said so far, is that the problem will be discussed with respect to consciousness' synchronic unity, as opposed to its unity over time.

[28] W. James, *A Pluralistic Universe: Hibbert Lectures at Manchester College on the Present Situation in Philosophy* (London-Bombay-Calcutta: Longmans, Green, and Co., 1909), 192.

[29] *Ibid*., 198–9.

[30] *Ibid*., 200.

sheer logical possibility, exist outside of the larger field of consciousness in which it actually occurs.[31] According to James, a person's psychical field at any one moment constitutes a non-decomposable unity; its several contents are to be viewed as 'aspects' *mutually determining and interpenetrating each other* rather than as 'parts' in any literal sense of this word. James' commitment to this view is implicit in his remark that 'each thought [a total moment of experience] is a fresh organic unity'.[32] The denomination 'organic unity' commonly refers to a totality whose parts are internally related, such that the nature of each essentially depends upon that of all others.

For the sake of illustration, imagine what it would be like to drink a cup of coffee in Naples as opposed to drinking it in Edinburgh: isn't it plausible to think that the different atmospheres of the two cities (the characteristically different colours, sounds, flavours *etc*. one experiences there) would make a difference to the coffee's taste? The two tastes, as they occur in the total states 'Coffee-in-Naples' and 'Coffee-in-Edinburgh', would seem to be different – *qualitatively*, and therefore also *numerically* – experiential occurrences.

Let's now consider the relation between the mind and the brain. It is unlikely that the panpsychist would want to reinstall the notion of the soul as an entity separate from the brain and causally interacting with it. Most likely, he is bound to speculate that the 'larger mind' of a human being is constituted by the 'lesser minds' of the neurons. This would seem to require that an experience belonging to (or even wholly constituting) a neuron's lesser mind be also felt by the larger subject which is the human mind. Such a view would involve what James refers to as 'the assumption that states of consciousness… can separate and combine themselves freely, and keep their own identity unchanged while forming parts of simultaneous fields of experience of wider scope'.[33] This underlying assumption can be recast as follows:

> THE SHARING PRINCIPLE – this is the view that an experience can simultaneously occur within two distinct psychical wholes – *i.e*. the very same experience can be felt by two different feelers, in this case, by the 'lesser mind' of the neuron and the 'larger mind' of the human being.

[31] D. Hume, *A Treatise of Human Nature*, edited by D. F. Norton and M. J. Norton (Oxford: Oxford University Press, 2000), 137–8.

[32] W. James, *The Principles of Psychology*, *op. cit*., 279n; see also 145 and 241.

[33] W. James, *A Pluralistic Universe*, *op. cit*., 181.

Once these notions are accepted, it becomes clear that the notion of mental combination involves a contradiction. If mental combination is to be possible, an experience *must* be felt by two different subjects while remaining numerically self-identical. But if the being of an experience is wholly exhausted by the way it feels, then an experience *cannot* be numerically the same while being felt by two different feelers. So, it would seem, mental composition is an impossible conception.

In order to see which options are actually open to a philosopher faced which this difficulty, it might be helpful to derive the contradiction in a more formal way. If mental combination is possible, then the same one experience, E, must be felt by the human mind M as well as by the neuronal mind N. At the same time, according to phenomenal essentialism, an experience is nothing but the way it feels; hence, the following two theses must *both* be true:

[1] $E = E$-as-felt-within-N
[2] $E = E$-as-felt-within-M

But how can this be? Since N and M are *per hypothesis* different, and M is surely going to have a richer content than N, this implies:

[3] E-as-felt-within-$N \neq E$-as-felt-within-M

By substitution of [1] and [2] in [3], one gets:

[4] $E \neq E$

But this is a clear violation of the logic of identity! In order for mental composition to be possible, an experience would then have to possess two natures – *it would have to be itself and yet another thing*. Surprisingly enough, James did not believe that the discovery of this contradiction amounted to a refutation of panpsychism; influenced by the then emerging philosophy of Henri Bergson, he chose to reject the logic of identity as adequate to an understanding of experience.[34] This is not an option to be ridiculed, but it is not to be welcomed either; ultimately, what one aims at in philosophy is rational, conceptual understanding.

Thus, it would seem that the panpsychist will have to renounce one of the principles involved in the derivation. Which one should he choose? The most likely candidate is the idea that experiences can be shared. This seems reasonable enough; after all, the privacy of our experience – *i.e.* the fact that each of us is the only individual who can possibly know what his or her experiences feel like – is a

[34] *Ibid.*, 212 and 214.

commonplace of every-day life.[35] Note that the point is not the *epistemological* one that we cannot *know* what another person's experience feels like, but *ontological*: experiences *are* such that they cannot be felt by more than one subject. And yet, to abandon the notion of a 'sharing' of experiences comes at a price; the panpsychist would seem to require this notion in order to make sense of mental composition and thereby account for the relation between the mind and the brain.

One final remark: it could be objected that the above argument does not show that there is any problem in principle with sharing across numerically distinct subjects if no other experiences were going on. This is true but, I think, irrelevant to the present case; it is only reasonable to assume that the human mind is a richer whole of experience than the (alleged) mind of a *single* neuron. Still, the question is important, as it suggests that there might be stronger ways of arguing for the privacy of our experiences – that is, without having to assume that there are differences in the subjects' other experienced contents and thereby without an appeal to the principle of phenomenal holism. As a matter of fact, a closer look at the principle reveals that it is not as evident as it might initially appear. James says little to support phenomenal holism, and while there are very obvious illustrations of it (a glass of wine has a better taste when enjoyed in a pleasant surrounding), the principle does seem somewhat implausible in other circumstances. Would the red of the book's cover in front of me have a different feel if I were not hearing music at this moment? It is difficult to tell, but there is room for scepticism here.[36]

7. Conclusion

Where does this leave us? Nagel has remarked that Cartesian dualism 'is usually adopted on the ground that it must be true, and often

[35] Using Samuel Alexander's little known but very apt terminology, 'knowing' our experiences means here apprehending them by way of 'enjoyment', simply by living through them as it were, rather than knowing them as one knows an object of contemplation. See his *Space, Time and Deity: The Gifford Lectures at Glasgow, 1916–1918*, Vol. I (London: MacMillan & Co., 1924), xiv.

[36] For a critical assessment of phenomenal holism, see B. Dainton, *Stream of Consciousness. Unity and Continuity in Conscious Experience* (London and New York: Routledge, 2000), 181–213 and see his paper in this volume.

rejected on the ground that it can't be true'.[37] The same could now be said of panpsychism, as the argument based upon the heterogeneity problem is counterbalanced by the problem of mental composition. Nagel's own conclusion is that 'panpsychism should be added to the list of… hopelessly unacceptable solutions to the mind-body problem',[38] but it would be premature to reject a doctrine supported by such a strong argument; rather, the hope is that what has been said might help to circumscribe the theory's 'logical space' – the legitimate area of argumentative manoeuvre for any attempt to make sense of mental composition.

If the above considerations are sound, then there are at least two requirements that any viable formulation of the theory must be capable of satisfying:

> [A] The panpsychist should reject the idea that larger experiential wholes are brought into existence by way of simple addition of lesser experiential realities; mental items must combine in a way that is categorically different from the way in which stones and bricks combine to form a house, or several pieces of a puzzle to yield the whole picture. These simply aren't the right combinatory models. (This might seem an obvious point to make; as it has been shown, however, it is also one that is easily overlooked.)
>
> [B] The panpsychist should fully acknowledge that experiences are private. This means that the notion of mental composition should not be construed as involving the idea that experiences can be literally 'shared' – 'owned' as it were by two different subjects. (To put it in idealistic terms: if the being of an experience is wholly exhausted by its *sentiri*, no experience can preserve its identity while felt by different feelers.)

Trivial as they might appear, these are severe constraints. Is there any *known* version of panpsychism that satisfies these demands? The only candidate that comes to mind is Leibniz's theory of causally independent monads. [A] The monads constituting our body do not sum up by way of addition to give rise to the human mind, which is a distinct, separate individual. And while he contends that our individual minds comprise an infinite number of unconscious *petites perceptions*,[39]

[37] T. Nagel, *The View from Nowhere*, *op. cit.*, 29.
[38] T. Nagel, 'Panpsychism', *op. cit.*, 193.
[39] G. W. Leibniz, 'Introduction to *New Essays on the Human Understanding*', in *The Monadology and Other Philosophical Writings*, edited and translated by R. Latta (Oxford: Clarendon Press, 1898), 370.

these are not to be thought of literally as 'parts' of our entire psychical whole; on the contrary, our minds are Leibniz's paradigmatic examples of genuine *units* or *indivisible* substances. [B] Furthermore, since monads are 'windowless', they do not literally share their contents with any other monad; although all monads perceive the *same* universe, each enjoys its own *private* mental world.

But wouldn't a recurrence to a Leibnizian metaphysics of windowless monads be a price too high even for a panpsychist to pay? Fortunately, this question can be postponed. In view of the strength of the panpsychist's case, what is needed is more work on the composition problem. In the passage quoted at the very beginning of this paper, Searle says that he cannot see 'any way that it [panpsychism] can cope with the problem of the unity of consciousness'. Undeniably, the problem is a difficult one; I have argued in this paper that – *perhaps* – there is a way. *We lack at present a firm understanding of phenomenal parts and wholes, but – pace Searle (and Goff) – it is equally true that we do not know that mental composition is impossible in principle*. Compared with the ferocious intensity with which other questions have been debated – such as, for example, mental causation or supervenience – this is an undeveloped area of research in contemporary philosophy of mind.[40]

University of Bern, Switzerland

[40] I would like to thank Leemon McHenry, Pauline Phemister and Galen Strawson for helpful comments on an earlier version of this paper.

Phenomenal Holism

BARRY DAINTON

Abstract
According to proponents of 'phenomenal holism', the intrinsic characteristics of the parts of unified conscious states are dependent to some degree on the characteristics of the wholes to which they belong. Although the doctrine can easily seem obscure or implausible, there are eminent philosophers who have defended it, amongst them Timothy Sprigge. In *Stream of Consciousness* (2000) I found Sprigge's case for phenomenal holism problematic on several counts; in this paper I re-assess some of these criticisms. Recent experimental work suggests cross-modal perceptual interference may be far more prevalent than expected. I argue that although these results do lend support to phenomenal holism in one of its guises, they do not support the strong form of holism espoused by Sprigge. I then move on to consider the relevance and impact of certain gestalt-related considerations, and argue that these considerations at best establish that the stronger form of holism applies to some parts of some experiential states, but not to all parts of all states, as Sprigge claims. I then consider a more promising way forward for anyone who wishes to defend an across-the-board holism of the strong variety, arguing that what is required is a form of phenomenal interdependence that is rooted solely in phenomenal unity. I conclude by outlining a case for thinking that an interdependence of this sort is a quite general feature of unified conscious states.

1. Phenomenal Interdependence

To make matters vivid and concrete, take a look at the shaded expanse shown in Figure 1 below. After focusing your attention on this for a few moments, reflect on the character of the visual experience you are now having, and consider this question: if some small part of this experience had been different over the past few seconds, would the other parts of your visual experience also have been different as a result of this?

To make matters still more vivid, we shall concern ourselves with just two small portions of your visual field: those corresponding to the A- and B-regions of E, as indicated by dotted lines in Figure 1 (we are now supposing that E represents the visual content which fills a part of your visual field). Now consider: if the B-region had been a subtly different shade, would your experience of the A-region have been different? Would A be affected in the slightest if the B-region were removed altogether? This experiment is easily

doi:10.1017/S135824611000007X

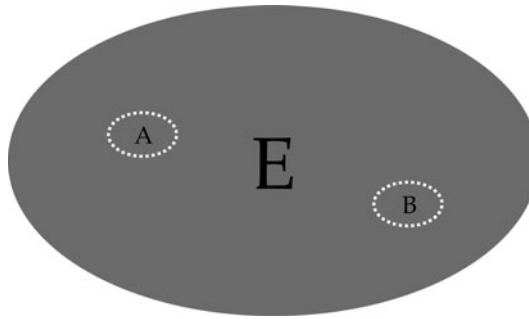

Figure 1.

performed: it suffices to cover the relevant part of Figure 1 with your finger, or a piece of paper. When you do this, does your experience of the A-region change in appearance in any discernible way? I strongly suspect that most of us would answer 'no' to each of these questions.

We have been considering the interdependence (or otherwise) of experiences (or parts of our experiences) within a single sensory modality, that of vision. What of cross-modal cases? Present purposes will be served by a single example. Focusing again on the precise phenomenal character of your experience of the B-region, would this character be in any way different if during this time you could also hear a faint buzzing noise? Or even a loud drill? The character of your *overall* state of consciousness would certainly be different – in one case you are experiencing the B-region without auditory accompaniment, in the other case you are experiencing it *with* an auditory accompaniment – but would the purely *visual* character of your experience be different in the two cases? Again, I suspect most of us would be inclined, quite strongly, to answer 'no'. This isn't to say the onset of the sound would produce no effects at all. If the noise were sufficiently loud, or unexpected, your attention might well be drawn away from your visual experience, and the character of your visual experience would certainly be altered as a result. Nonetheless, were this to occur, the visual and auditory contents that you are apprehending would still remain distinct. In our ordinary experience at least, there is no discernible *fusion* or *interpenetration* of phenomenal qualities belonging to different sense-modes: sounds remain purely and exclusively auditory in nature, visual contents remain purely and exclusively visual in nature.

These *prima facie* plausible claims about the independence of the intrinsic characteristics of phenomenal contents are quite mistaken

if the doctrine I shall here call 'phenomenal interdependence' is true. The doctrine can be formulated in different ways – and indeed, it comes in importantly different forms, as we shall see – but the rough-and-ready formulation below will suffice to convey the basic idea:

> If X and Y are phenomenal contents that are apprehended together, in a single state of consciousness by a subject *S* at a time *t*, then the intrinsic character of each of these contents impacts in a significant way on the character of the other.

Now the doctrine of phenomenal interdependence is by no means obviously true, and we have just encountered an example of why: in the case of the visual field which results from looking at Figure 1, regions A and B are phenomenal contents that are experienced together within a single state of consciousness, but it seems quite plausible to suppose that the intrinsic phenomenal qualities of A would not have been significantly different had A been experienced in the absence of B. That said, if the phenomenal interdependence doctrine *were* true, it would have very significant implications for our understanding of the nature of conscious experience, and conscious states. For this reason alone the doctrine should be of interest to those concerned with understanding the nature of consciousness.

2. Sprigge's Phenomenal Holism

The phenomenal interdependence doctrine may not be obviously true, but it does have its supporters. In recent years, one of its most eminent champions has been Timothy Sprigge. He defends the doctrine at some length in his own distinctive terms and manner in *The Vindication of Absolute Idealism*, where it plays a key role in his idealistic metaphysic – see in particular Chapter 5, part 3. But although Sprigge recognised the importance of the doctrine, he did not take himself to be propounding anything novel, far from it. The notion that the parts of a conscious state are profoundly interdependent may sound somewhat odd to contemporary ears, but it was comparatively commonplace among nineteenth century philosophers. In his *James and Bradley*, Sprigge points out that although they differed on much else, both Bradley and James subscribed to the view that our states of consciousness 'at any one moment are wholes such that every element within them is so coloured by the totality

that they could not exist again without difference in another state of consciousness'.[1] Sprigge himself subscribed to this view:

> A holistic relation is strong if the kind of whole its terms unite in forming has a character which so suffuses its every element that no element with some difference from it in character could be found without a whole of just that sort.... All holistic relations between terms actually given in experience appear to be strongly so.[2]

As I mentioned earlier, the phenomenal interdependence doctrine can be formulated in different ways, and it comes in a variety of stronger and weaker forms. A proponent of *complete* phenomenal interdependence claims that all parts of all unified states of consciousness are holistically interrelated, whereas the proponent of the weaker *partial* phenomenal interdependence maintains only that *some* parts of unified conscious states are so interrelated (or alternatively: only some parts of *some* states). It is also possible to hold that phenomenal interdependence applies necessarily, or only contingently. If the latter obtains, states that are actually interdependent might not have been, if the former obtains, states that are actually interdependent could not possibly be otherwise.

Sprigge subscribes to the interdependence thesis in its strongest guise: complete and necessary. Given this, it is reasonable to conclude that he believes that the parts of conscious states are *by their very nature* holistically interrelated. But why believe this to be true? In the *Vindication* he motivates this doctrine thus:

> That some holistic relations are strongly so is readily revealed. Consider the character of a painting and the relation between its parts, when the painting is seen as a whole. Consider the painting, that is, as a total presence in someone's perceptual field (where else?). It is a commonplace of aesthetics and of right-minded psychology, but something we can each discover for ourselves, that every detail in the painting as a complete presentation has some difference, even within its own bounds, from what the detail would have if it were seen apart, or in another whole (as can be arranged on various physical bases), although if when one sees the detail 'on its own', or in a different whole, one has a lingering memory of the whole, it will be less different

[1] T. L. S. Sprigge, *James & Bradley: American Truth and British Reality* (Illinois: Open Court, 1993), 2.

[2] T. L. S. Sprigge, *The Vindication of Absolute Idealism* (Edinburgh: Edinburgh University Press, 1983), 218–9.

> from its previous self than otherwise. An eye, as it figures in a certain painted face, will supply a good example. Certainly, the same identical shape and pattern of colour can be present in a different whole, but one cannot think of what lies within the eye's own bounds as having a character unaffected by the whole it helps to form. Indeed, the mere place at any moment of any visual phenomenon in the visual field, high or low, to the left or the right, gives it a different character. So far as it makes sense to think of it as a particular component in experience, having its individual character within its bounds, that character is affected by the whole, though one may prefer to say that it really has no such separate character.[3]

Sprigge here suggests that phenomenal interdependence, at least in certain cases, is clearly manifest in our own experience, provided we pay sufficiently close and careful attention. But he also maintains that the same sort of interdependence applies across the board, to *all* parts of unified states of consciousness. Following Sprigge, let us use the term 'total state' or 'total experience' to refer to our overall states of consciousness at any one time (*i.e.* a momentary or very brief temporal cross-section of our streams of consciousness). It may seem that our total states of consciousness include parts that might conceivably exist in a total state of a different overall character, but this is wrong:

> These, however, are in an important sense not real parts of the total experience, nor are any of its other components. This is because they lack an individual essence which could be specified or grasped without reference to the whole to which they belong. That is, an attempt fully to grasp what one of these components is, within the limits of its own being so to speak, could not specify a determinate possible form of being it actualises, since its fully determinate form of being is something which involves its particular sort of contribution to the character of the whole.[4]

This fully general claim is a very strong one indeed. If it is true, it has significant implications for the nature of conscious states. But how convincing are the considerations which Sprigge brings to bear in attempting to establish that it *is* true?[5]

[3] *Ibid*., 219.
[4] *Ibid*., 170.
[5] As will already be clear, I am focusing here on the experienced-based (or phenomenological) case for holism which Sprigge develops. It is important to note – and I am grateful to Pierfrancesco Basile for reminding me – that Sprigge also subscribes to (a form of) holism on more general

3. Cross-Modal Interference

When discussing these issues in my *Stream of Consciousness*, I argued that the sort of consideration to which Sprigge appealed fell a long way short of establishing the conclusion he sought.[6] I did not deny that there are *some* localised phenomenal interdependencies. Indeed, much of what Sprigge says about the effects of changing the context of an eye is very plausible. As can be seen from Figure 2, an eye of a dog that is set in the surroundings of a (friendly-looking) dog looks rather different from the same image placed in a different context.

These context-induced changes are quite subtle, and not easy to describe, but they are nonetheless real for that. Also, importantly, it seems right to ascribe them – in part at least – to changes which affect the intrinsic visual qualities of the experiences in question: the eye on the right *alters in visual appearance* when the surrounding dog is eliminated. Other examples of localised intra-modal

---

metaphysical grounds: 'I do not see how one can ever imagine two or more things related to one another in any way one likes to specify, without this being an imagining of them as forming a specific sort of whole together, or joining with other things in doing so' (*Vindication*, *op. cit.*, 208). In short, Sprigge was of the view that we cannot coherently view two items – of *any* kind – as being connected by a genuine relation unless they form parts of a single whole. However, the doctrine that real relations can only hold between the parts of a whole does not, in itself, entail the stronger holistic doctrine (more clearly distinguished in section 4 above) that the intrinsic characteristics of the parts of a whole are influenced, in a distinctive way, by the character of the whole to which they belong. So far as I can see, Sprigge's grounds for subscribing to this stronger form of holism are phenomenological. I should also add that, for Sprigge, the significance of these holistic considerations is not confined to our understanding of the nature of consciousness; it extends to the whole of reality: 'I think that real relations between things can only be conceived as the way in which they join together to form a whole, and that the whole which a congeries of experiences can form is itself an experience . . . So all the experiences which fill up the world must ultimately join together as part of one great Cosmic Experience' (T. L. S. Sprigge, 'My Philosophy and Some Defence of It' in *Consciousness, Reality and Value*, edited by P. Basile and L. McHenry (Ontos Verlag: Heusenstamm, 2007), 301). Interesting though it is, I will not be concerned with Sprigge's case for absolute idealism in what follows.

[6] B. Dainton, *Stream of Consciousness* (London: Routledge, 2000; expanded 2nd edition, 2006).

Figure 2.

interdependencies of this sort are not difficult to find. In the case of the well-known Müller-Lyer illusion, the direction of the 'tail-fins' makes a difference to the apparent length of the two horizontal parallels. But, I suggested, while it is right to recognise that such interdependencies exist, it is questionable whether they are enough to sustain Sprigge's very strong conclusion.

First of all, these interdependencies look to be contingent rather than necessary. The Müller-Lyer effect is certainly due to the idiosyncrasies of the human visual system. As for the case depicted in Figure 2, it is not implausible to suppose that a visually-endowed subject with no prior acquaintance with either dogs or the appearance of mammalian eyes, would see both eyes in much the same way. If the effects are only contingent, we have no reason to think that phenomenal interdependencies are rooted in the very nature of conscious experience.

Secondly, from the fact that *some* parts of some total experiential states are interdependent, we cannot automatically, or without further argument, conclude that the same applies to *all* parts of *all* such states. Our earlier example is relevant here. I suggested in connection with Figure 1 that it is not obvious that the intrinsic *visual* qualities of the B-region would be in any way affected if the contents of the A-region were altered or absent. The inter-modal case looks to be even less promising from the point of view of the would-be holist: how plausible is it to suppose that the visual quality of either region be altered in the slightest if these visual contents were accompanied by an auditory sensation? In *Stream of Consciousness* I laid considerable emphasis on the apparent scarcity of intermodal interdependencies:

> It is hard to believe that one's current auditory experience is significantly responsive to small variations in one's visual

> experience, or that the character of one's current tactile experience would be different were one's current olfactory experience other than it is. Quite generally, at any given moment, any pair of co-conscious experiences belonging to different perceptual modalities seems largely – and typically completely – unaffected by each other..... If the character of experiences in different sensory modalities are generally independent, which contrary to Sprigge's claim they seem to be [then] complete holism cannot be grounded in phenomenal interdependence.[7]

Suppose, just for a moment, that these claims are correct. If the bulk of our experience is *not* phenomenally interdependent, why did Sprigge think otherwise? I outlined two possibilities. The first is that he had succumbed to 'consciousness mysticism', a condition of intellectual intoxication produced (or so I speculated) by spending prolonged periods engaged in intense introspection, and which leads its victims to ascribe exaggerated or chimerical properties to consciousness.[8] But I went on to mention a second and more charitable possibility: he may simply have exaggerated the extent to which the kind of interdependency described above actually obtains. It is not difficult to see how this might have come about.

It is very plausible to think that phenomenal unity – the sort of unity we find in our conscious states at any given time – is a unity of a distinctive kind. If this is so, it seems equally reasonable to suppose that experiences that are unified in this way might be affected in some way as a result. Since *all* parts of a total conscious state are phenomenally unified, the phenomenal effects (as we might call them) generated by this mode of unity will also extend to all parts of these states: the resulting holism will be complete, rather than partial, and necessary, as opposed to contingent. The only remaining question is the *kind* of effect or change that we are dealing with here. Sprigge was, of course, well-acquainted with the doctrine to which both Bradley and James subscribed, that every element in an experiential whole is invariably 'coloured' by the totality it finds itself in, to such an extent that the part could not exist in isolation, or in a whole of a different kind. Finding some confirming evidence for this contention in his own experience – see the quotations above – he not unreasonably concluded that this sort of interdependency obtains across the board: that it applies to *all* experiences, even if this is not obvious.

[7] *Ibid*., 194–5.
[8] *Ibid*., 195.

This unity-related line of reasoning may have its merits, but as we have seen, Sprigge's conclusion can seem highly dubious. If we confine ourselves to what is revealed by introspection, it simply does not seem to be the case that the intrinsic visual characteristics of simple expanses of colour are 'coloured' or influenced in any introspectively discernible way by accompanying experiences in other sensory modalities, or by accompanying mental images or conscious thoughts. If the conclusion is false, we have no option but to conclude that Sprigge's reasoning is unsound. As for where it goes astray, the most obvious candidate is the extrapolative step: from the fact that *some* parts of our total states of consciousness impact on the character of other parts, it does not follow that they *all* behave in this manner.

4. Sprigge Vindicated?

We will be returning to the issue of whether it is possible to mount a plausible case for a generalised or *complete* inter-experiential holism. But before moving on I want to take a step back. When I confidently pronounced in *Stream of Consciousness* that inter-modal interdependencies were non-existent, or at least very rare, was I in fact correct? A growing body of empirical results from psychology and psychophysics suggests that I may well have been quite wrong.

That there are *some* instances of inter-modal interference has been known for some time. The 'ventriloquist illusion' is perhaps the most familiar. Even though we know that the ventriloquist's doll isn't really saying anything – all the sounds are actually emerging from the (misleadingly motionless) mouth of the ventriloquist – it seems as though the words are being produced by the doll, whose lips *are* moving. In this case, visual data influences the apparent location of sounds. The 'McGurk effect' is another well-established instance of audio-visual interdependence, one which also involves speech-perception. But in this case, rather than visual data influencing the apparent location of sounds, the auditory and visual influence each another. Our lips move in different ways when we utter different phonemes – *e.g.* when making the *mmmm* sound our lips come together, when making the *nnnn* sound they don't. When shown a video film featuring a close-up of a subject whose lips make the movement associated with the *gaa* sound, but who in fact produces a *baaa* sound, what most of us hear is a *daaa*. It seems that in the case of speech perception at least, if our brains are presented with auditory and visual information which conflicts, they seek a compromise solution. The effect is robust – it usually persists even when one

knows the trick being worked – and is not confined to native speakers of English.[9]

These two effects may be the best-known examples of 'cross-modal' interference, but they are not alone. Recent investigations have uncovered a variety of other cross-modal interactions, involving different sensory modalities, in different combinations:

- *the sound-induced flash*: when a single flash of light is accompanied by several auditory 'beeps', subjects tend to perceive several flashes of light, rather than just one.[10]
- *the touch-induced flash*: if subjects are shown a single flash accompanied by two taps on the skin, they tend to see two flashes.[11]
- *the parchment skin illusion*: when subjects are asked to rub their hands together while listening to high frequency sounds delivered via headphones, they report that their skin feels unusually smooth and dry (like parchment); if the high frequencies are dampened, subjects report that their hands feel unusually smooth and moist.[12]
- *sound-induced changes to perceived crispness and fizziness*: the apparent 'crispyness' of a potato-crisp depends on the sounds heard while munching on it – damp down the high-frequencies and it will seem soft and stale; in a similar vein, fizzy water on the tongue feels fizzier when accompanied by high-frequency sounds, and electric toothbrushes feel smoother in the absence of high-frequency sounds.[13]
- *motion-after-effects transferring between sight and touch*: in the well-known 'waterfall' illusion, if you stare for some time at

[9] See H. McGurk and J. MacDonald, 'Hearing Lips and Seeing Voices', *Nature* **264** (1976), 746–8. There are a good many examples of the effect readily available on the web (including several on *Youtube*), some more effective than others.

[10] L. Shams, Y. Kamitani and S. Shimojo, 'What You See is What You Hear', *Nature* **408** (2000), 788.

[11] A. Violentyev, S. Shimojo and L. Shams, 'Touch-induced Visual Illusion', *Neuroreport* **16:10** (2005), 1107–1110.

[12] V. Jousmaki and R. Hari, 'Parchment-skin Illusion: Sound-based Touch', *Current Biology* **8:6** (2006), 190–191.

[13] M. Zampini, S. Guest and C. Spence, 'The Role of Auditory Cues in Modulating the Perception of Electric Toothbrushes', *Journal of Dental Research* **82:11** (2003), 929–32. Also, M. Zampini and C. Spence, 'The Role of Auditory Cues in Modulating the Perceived Crispness and Staleness of Potato Chips', *Journal of Sensory Studies* **19:5** (2009), 347–63.

> downward-moving water, and then focus instead on the neighbouring rocks and trees, you will see the latter seemingly move upward, even though in reality they are motionless. Surprisingly, a similar effect can be induced via the sense of touch. Subjects are asked to spend a few moments staring at a screen filled with motionless horizontal stripes, after which time the palm of their hand is stimulated (by electronically controlled pins) so as to provide them with the impression that an object is sweeping up or down over their skin; the subjects then report that the lines on the screen have started moving, in the opposite direction to the motion they feel on their skin. The effect works in reverse: motionless pins seem to start moving when subjects are observing horizontal stripes moving up or down on the screen in front of them.[14]

These examples of cross-modal interference are intriguing in their own right, but they also potentially have significant implications of a more general kind. Vision has often assumed to be the dominant sensory modality, and it has been thought that the processing of visual information is independent of goings-on in other modalities. The sound- and touch-induced flash illusions cast a large shadow over these assumptions. They demonstrate that in the case of conflicting or ambiguous stimuli, vision does not invariably trump the other senses, and in some circumstances (at least), the processing of visual information is *not* independent of the processing of data in other sensory modes. Although much remains to be investigated and discovered, these various results are pointing in the same general direction: it may well be the case that none of our senses are independent of any of the others. In ordinary circumstances, confronted with the task of providing us with perceptual experience which corresponds with our external environments, on the basis of flimsy, fleeting and often conflicting sensory data, our perceptual systems are only too willing to allow information deriving from some modalities to override others. In this quest for consistency and coherence, no sensory modality is immune to potential interference of this kind.[15]

[14] T. Konkle, Q. Wang, V. Hayward and C. Moore, 'Motion Aftereffects Transfer between Touch and Vision', *Current Biology*, DOI: 10.1016/j.cub.2009.03.035 (2009), http://dx.doi.org/10.1016/j.cub.2009.03.035.

[15] For further useful discussion of these matters, see C. O'Callaghan, 'Seeing What You Hear: Cross-Modal Illusions and Perception', *Philosophical Issues* **18:1** (2008), 316–38.

If this is right, what are the implications for our current concerns? One thing is very clear: I was wrong when I claimed that our sensory modalities are almost completely independent, and hence that the character of our experience in one modality at a given time *t* would almost certainly be exactly the same if the character of our experience in other sensory modalities were different at *t*. However, my being mistaken in this respect does not in itself mean that phenomenal holism, in the form espoused by Sprigge, is correct. And this for several reasons.

The phenomenal interdependence doctrine comes in stronger and weaker forms, and as noted earlier, Sprigge subscribes to the doctrine in its strongest form: he holds (i) that *all* parts of *all* total experiences are interdependent, and (ii) that this obtains as a matter of necessity. The fact that cross-modal interference, of the kind we have just been looking at, is more prevalent than has sometimes been thought does not, in itself, establish strong, Sprigge-grade holism. Since 'more prevalent' does not mean 'extends to *all* parts of *all* total states of consciousness', we are still looking for a reason for supposing that Sprigge was right to opt for the *complete* interdependence doctrine. Furthermore, since there is no reason to think these interference effects are due to anything more than the peculiarities of human (or mammalian) sensory systems, there is no reason to think these interference effects are anything other than contingent.

A second point is epistemological, and relates to the grounds we have for accepting a phenomenal interdependence claim, of whatever strength. Sprigge's case for holism rests on phenomenological considerations: the evidence which suffices to establish that interdependence obtains is available to introspection, or so he suggests. Irrespective of whether he is right about this, most of the interference effects outlined above are invisible to introspection. Indeed, that these interdependencies exist at all only emerges under unusual experimental conditions. And of course, since they come as a complete surprise to most of us, it is reasonable to conclude that nothing in our everyday experience suggests they exist.

A third point relates to the *kind* of interdependence that is at issue. While there may well be a greater quantity of inter-modal perceptual interference than one might have supposed, prior to learning of the experimental results from psychology, it is doubtful whether these interdependencies are of the kind which interested Sprigge. To bring this out, it will help to make explicit a further distinction. Let us suppose we have a total state of consciousness *S* at a time *t*, which can be divided into proper (experiential) parts, P_1, P_2, P_3 …

P_N. We can now distinguish two ways in which these parts can be holistically interrelated:

> *Deep Interdependence*: in actual fact, all of P_1, P_2, P_3 … P_N exist and form part of *S* at *t*; each of these experiential parts has one or more intrinsic phenomenal features which (in some manner) reflect(s) the character of the whole, in such a way that none of the parts could exist in a total state of consciousness with an overall phenomenal character different from that which *S* possesses at *t*.
>
> *Shallow Interdependence*: in actual fact, all of P_1, P_2, P_3 … P_N exist and form part of *S* at *t*. If any of these experiential parts had been absent, or replaced by a part with a different phenomenal character, then the phenomenal character of some or all of the remainder of *S* at *t* would have been different as a result. Despite this, *S*'s parts are not essentially bound to a whole of this particular type: each of P_1, P_2, P_3 … P_N could exist in a total state whose character differs from that of *S* at *t*.

These formulations are deliberately vague – I have, for example, made no attempt to say anything about the precise manner in which experiential wholes impact upon the parts in the case of Deep Interdependence, hence the ' … (in some manner) …' but for our immediate purposes they will serve. If a collection of experiential parts are Deeply Interdependent, then the whole to which they belong impacts in a distinctive way on the phenomenal character of each of the parts, with the result that none of the parts could exist in a whole of a different overall type. (Here the standard assumption that the precise phenomenal character of a token experience is essential to it is in play.) In contrast, where only Shallow Interdependence obtains, it is also the case that if some parts of an experiential whole were absent, or replaced by a substitute with a different character, then the other parts would not have the character they actually do have, but in the absence of any 'imprinting' of the character of an experiential whole onto its parts, there is nothing to prevent experiences with the character of these parts existing in experiential wholes of a different type.

Now, it is clear that Sprigge subscribed to the Deep Interdependence doctrine. In the passages cited earlier he says that all holistic relations between elements given in experience are of the *strong* variety, and that a 'holistic relation is strong if the kind of whole its terms unite in forming has a character which so suffuses its every element that no element with some difference from it in

character could be found without a whole of just that sort'.[16] It is natural to read Sprigge as holding that Deep Interdependence arises as a consequence of the distinctive way in which experiential parts are related within total states of consciousness. Irrespective of whether he is right about this, now that we have the distinction between the two modes of phenomenal interdependence clearly in view, the question we need to consider is the following: is there any reason to think that the intermodal interference effects outlined earlier involve Deep rather than merely Shallow Interdependence? I cannot see that there is. In the case of the sound-induced flash, for example, the presence of a second 'beep' causes subjects to see an illusory second flash, but this flash is a perfectly ordinary visual experience: there is no fusion or interpenetration of beep-content and flash-content, and it would be perfectly possible to experience a phenomenally indistinguishable flash of light *in the absence of any beeping sound*. Much the same applies in the parchment skin case: thanks to the presence of high-frequency sounds, one's hands feel rougher to the touch than would otherwise be the case, but the auditory and tactile sensations remain (seemingly) entirely un-merged or un-fused. And it is very plausible to think that exactly similar tactile sensations could have existed in total states of consciousness of a different kind – states which do *not* include high-frequency sounds, even if the skin on one's hands might need to be significantly rougher to provide these sensations, if the sounds were absent. If this is right – and it does seem plausible – then these interference effects are manifestations of Shallow rather than Deep Interdependence.

Pulling these points together, the conclusion is clear: cross-modal interactions of the kind we have been considering may be a good deal more common than has usually been thought, but they do not provide Sprigge with what he needs.

5. Gestalt-Based Holism

If the parts of our total states of consciousness are interdependent in the deep way which Sprigge, following here in the footsteps of Bradley and James, believed them to be, we will need to look beyond the cross-modal perceptual interference effects we have been considering latterly. But where? Do we have any reason to believe experiential parts *can ever* be Deeply Interdependent? Sprigge's example of the eye-in-a-painting may provide some support for the

[16] T. L. S. Sprigge, *Vindication*, *op. cit.*, 218.

Figure 3.

possibility of Deep Interdependence, but there are simpler and (somewhat) more straightforward cases that point in a similar direction. I will confine my attention here to these simpler cases.

If you stare for a minute or so at the vertical lines depicted in Figure 3, you will probably see them undergo a series of 'aspect-shifts': for a few seconds you will see them as forming two groups of three, then three groups of two, then back to two groups of three, and then (perhaps) as a group of four with a single outlier to the right and to the left. As the lines form one 'perceived whole' after another, their appearance undergoes a subtle alternation: the line third in from the left *looks* somehow different when it is being seen as part of two groups of three. Do we have here a case of Deep Interdependence, a case where the intrinsic phenomenal character of the part is influenced by the kind of experiential whole to which it is perceived as belonging? Certainly, some Gestalt theorists seemed to take this view, when they maintained that 'structured' or 'organized' experiential wholes exert an influence on the character of their component parts. Here is one such:

> Since [sensory] data exhibit phenomenal features only derived from the configuration into which they are integrated, it follows that *such a configuration cannot be considered as built up out of the parts* … if a constituent of a configuration is isolated and taken by itself as an independent and self-contained element, it may be affected so radically and by such deep reaching modifications as to destroy its phenomenal or experiential identity, the constancy of the external stimuli notwithstanding.[17]

I discussed this and similar claims in *Stream of Consciousness*[18], and expressed some scepticism. Yes, the lines in Figure 3 do look different when they are perceived as belonging to different configurations – and

[17] A. Gurwitsch, *Field of Consciousness* (Pittsburgh: Duquesne University Press, 1964), 114; italics in the original.

[18] *Op. cit.*, §§ 8.5 and 8.6.

the same applies in cases featuring other kinds of perceived whole – but we must remain wary. It is one thing for a line to look subtly different when seen as belonging to a group of three rather than a group of two, but to assess whether this difference amounts to an instance of Deep Interdependence we need to consider questions of the following sort. When (for example) the line in question is seen as the centre-most part of a group of three, does it have an appearance which a similarly shaped line could *only* have when perceived in such a configuration? Or could a line with exactly similar intrinsic phenomenal features exist all by itself, or in a group of a different kind? The latter proposal, I suggested, seems the more plausible. A little experimentation suggests that the intrinsic visual features of the line in question *are* discernibly different when it is perceived in different configurations, but these differences are of an unremarkable kind: the line may appear to vary slightly in thickness, or the distance between the lines may seem to change in some small way. Accordingly, when the same (physical) line is perceived as belonging to a group of three it may have an appearance that differs slightly from what it would have if it were to appear as a part of a group of two, but there is nothing to prevent a line with precisely the same intrinsic features appearing on its own, or as a part of a different combination (*e.g.* as a part of a group of four), even if it is likely that different physical stimuli would be needed to generate the same phenomenal appearances in these cases (*e.g.* the line on the page would have to be slightly thicker, or thinner or darker). If this is right, then it seems that gestalts do not, after all, present us with instances of Deep Interdependence.

While much of this still strikes me as plausible, it may well be that gestalts have properties that I failed to notice or address in my earlier discussion, and which may be of assistance to the holist. Consider that familiar illustration of an aspect-shift, the duck-rabbit, as shown in Figure 4.

As you see the figure take on the aspect 'duck' then 'rabbit', then 'duck' again, its general appearance differs in quite a dramatic way: when it looks like a rabbit, it *looks* very different from how it looks when seen as a duck, even though the visual stimuli, in the form of the markings on the page (or screen) are precisely the same. So far, so familiar, but now focus your attention onto just *part* of the picture, *e.g.* the ears and/or beak region, as picked out in Figure 5.

The question we need to consider is whether this part of the visual whole also has a different visual appearance when the whole appears under the aspect 'rabbit' than it does when it appears under the aspect 'duck'. The answer, I take it, is plain: there is a difference, and the difference is a significant one. When the whole figure looks like a

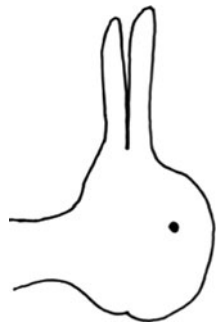

Figure 4.

duck, the encircled region looks like a *beak*, whereas when the figure looks like a rabbit, this same region looks like a *pair of ears*. It is clear that in this case at least, the visual alterations which accompany aspect change extend to the parts of the whole that is perceived. Drawing on this, the holist can argue as follows:

> The duck-rabbit is a simple illustration of a more general phenomenon. What it illustrates, in so very striking a manner, is the fact that there is more to visual experience than colour-patterns of one kind or another. The additional ingredient is *meaning*, or if you prefer, *representational content*. For the switching or 'dawning' of aspects surely is best explained in these terms: the same few lines on the page are seen *as a duck*, and then *as a rabbit*. The fact that high-level content of this kind is present *in* sensory experience – and not, as some have argued, wholly a matter of accompanying judgments or beliefs – is demonstrated by the way the duck-rabbit changes in visual appearance when

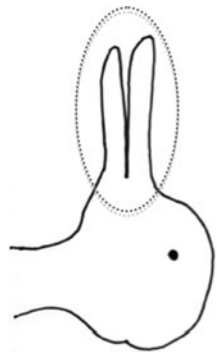

Figure 5.

> it is seen under different aspects. This fact supplies plausible and compelling grounds for recognising that Deep Interdependence is a reality. Evidently, the reason why the encircled part of Figure 5 looks like a pair of ears (when it does) is because the other parts of the figure are jointly suggestive of a rabbit. Hence we are dealing here with a case in which the character of the whole influences the character of the parts. Importantly, in this sort of case, the influence is of a highly distinctive sort, for the character with which the whole imbues the part is *not* of a kind which could or would exist in a whole of a different kind. After all, we do not see the rabbit-ears attached to a duck, or the duck-beak attached to a rabbit. It is the non-transferability of meaning-imbued parts which generates phenomenal interdependence of the strong variety in such cases.

This general line of argument points to a number of interesting avenues that are worthy of further exploration, but I am doubtful that it will supply Sprigge with what he needs.

For the argument to be viable, it has to be the case that high-level concept-laden representative content can be present in sensory experience. A case can certainly be made for thinking that when we see a duck *as* a duck (or a door *as* a door or hear a duck's quacking *as* a duck's quacking) the content 'duck' (or 'door' or 'quacking') is as much a part of our sensory experience as any colour or sound qualities. But a case can also be made for thinking that this is not the case. Couldn't a subject – perhaps an animal or young infant – look at a door, and have an experience which in all *purely visual* respects is indistinguishable from the experience we ourselves have when we look at a door, but without their experience being imbued with anything resembling the concept 'door'? It is not obvious that this is wrong. The question of whether and in what form anything resembling conceptual or representational content features in our perceptual experience itself is an important one, and as yet very much unresolved. A recent issue of the *Philosophical Quarterly* was largely given over to this topic, and whereas some authors – *e.g.* Bayne and Siegel – defend the view that the phenomenal features of perceptual experience can include higher-level categorical or kind-properties (*e.g.* … *is a tomato* or … *is a pine tree*), others – Byrne, Pautz and Price – find the claim that the contents go beyond basic sensory properties (shape, size, colour *etc.*) highly problematic.[19]

[19] See T. Bayne, 'Perception and the Reach of Phenomenal Content', *The Philosophical Quarterly* **59** (2009), 385–404; S. Siegel, 'The visual

To cut a long story short, holists who rest their case on the doctrine that our sensory experience is imbued with high-level content are offering a sizeable hostage to fortune.

Setting this point aside, even if we suppose that sensory experience does have the sort of content the holist needs, it still may well be not enough. First of all, in the case of objects perceived under a certain aspect, is the character of the parts of the resulting phenomenal wholes really such that it is impossible for the parts to exist in wholes of a different type? So far as the duck-rabbit is concerned, I find that – on some occasions, if not all – I can succeed (after a little effort) in seeing the figure as embodying both aspects simultaneously: I see the figure as a rabbit with a duck's beak emerging from the top of its head, or alternatively, as a duck with a rabbit's face for a head. A few informal tests carried out on other subjects suggest I am not alone in this. (If when performing the experiment you find that you only succeed in seeing a rapid alternation of duck

---

experience of causation', *The Philosophical Quarterly* **59** (2009), 519–40; A. Byrne, 'Experience and Content', *The Philosophical Quarterly* **59** (2009), 429–51; A. Pautz, 'What are the Contents of Experiences?', *The Philosophical Quarterly* **59** (2009), 483–507; R. Price, 'Aspect-switching and Visual Phenomenal Character', *The Philosophical Quarterly* **59** (2009), 508–18. Needless to say, the issue is by no means a new one, and it divided Gestalt theorists themselves: whereas some of these theorists held that structural or aspectual features are present in our base-level sensory experience, others held that these features are to be found only in higher-level conscious acts or qualities. In a useful chapter on this topic, Smith summarises thus: 'The fact that our experience is structured, is, according to Ehrenfels, a matter of certain special *Gestalt qualities* which are given in special experiences, superadded to our experiences of sensory elements. A two-level theory of this sort was … characteristic of that "Austrian" approach to complex experience which was developed by Ehrenfels, Meinong, Witasek, Benussi, Bühler, and their followers. According to the later "Berlin" approach [of Wertheimer, Koffka and Köhler], in contrast, a collection of data (or any other psychological formation) does not *have* a Gestalt on a second level. Rather, it *is* a Gestalt, a whole whose parts are themselves determined as being such that they can only exist as parts of a whole of this given kind. The significance of this distinction, or of the transition from the Austrian theory of Gestalt as *quality* to the Berlin theory of Gestalt as *whole*, cannot be overestimated' (B. Smith, *Austrian Philosophy* (Illinois: Open Court, 1994), 245). In *Stream of Consciousness*, I was reluctant to extend high-level content or conceptual content to perceptual experience, but did allow that meaning is certainly to be found in our perception of speech and writing (*Stream of Consciousness*, *op. cit.*, §8.7).

and rabbit, try telling yourself first what you are looking at is a *duck-rabbit*, the unfortunate upshot of an experiment to fuse ducks and rabbits at the genetic level.) Not all instances of aspect-switching permit this permutation and re-combination of parts. The 'staircase illusion', where the same drawing can suddenly flip from a staircase running in one direction to a staircase running in a different direction, is one such. But even a few counter-instances will be enough to worry anyone seeking to develop a case for *complete*, rather than partial, phenomenal holism on gestalt-related factors.

A further and probably stronger consideration points in the same direction. Suppose Sprigge is right, and holism extends to all parts of every unified state of consciousness. If this holism is to rest on gestalt-related considerations, then *every* combination of the experiential parts which jointly compose such states must form gestalts. No doubt many do: it seems plausible to suppose that gestalts feature prominently in our ordinary visual experience. When I look out of the window, I see a road, some trees, a car, a woman walking a dog: here the various parts of my experience corresponding to a familiar stretch of the *road* constitute a gestalt, as do parts corresponding to the *trees*, and also the *cars*, likewise the *woman walking a dog*. Moreover, given that this scene – or much of it – is a very familiar one, it may well be that the entire combination of road + car + trees + woman-walking-dog *also* constitutes a gestalt. But other everyday examples are less promising in this regard. If I walk into town, some parts of my visual experience will form gestalts – a familiar row of buildings, a street-corner of long acquaintance – but there are many others that will not: the combination of the familiar street-corner with the person now standing there (whom I have never before seen) and the vehicle parked there (likewise entirely new to me). More generally, although many of the individual *objects* which feature in our visual experience are gestalts, a great many of these objects *when taken together* do not form gestalts, simply because these particular combinations of elements are new to us, or unique, or both.

The situation worsens when inter-modal combinations are taken into account. Earlier today, while gazing at the familiar scene visible from my window, I entertained a mental image of the Eiffel tower, enjoyed the flavour of cheese from my lunchtime sandwich, and simultaneously felt a tingle in my toe. The elements of my visual field may form a gestalt, but does the combination of cheese-related gustatory sensation, the toe-tingle and the mental image of the Eiffel tower? There is no reason to think so. These very diverse experiential elements do not form a pattern of any recognisable

kind; taken together, they lack anything which could plausibly be called organisation or structure: in these respects they are quite *un*like the patterns of dots on the face of a die, or the sequences of notes which comprise familiar melodies. It may well be true that any combination of heterogeneous experiences *can* form a gestalt. Someone who, for whatever reason, takes a serious interest in the experiential combination 'mental image of Eiffel tower + tingle in the toe' would no doubt feel a sense of familiarity and recognition whenever they enjoyed this pairing of experiences: for *this* subject these experiences probably would form a gestalt. But obviously, this sort of case is very much the exception, rather than the rule: for most subjects, most heterogeneous combinations of experiences do not form gestalts. Since it seems plausible to suppose that substantial parts of our ordinary streams of consciousness are composed of precisely this sort of experiential combination, the prospects for building a compelling case for *complete* phenomenal holism based on gestalt-related considerations seem dim.

6. From Unity to Holism

If gestalt-related considerations will not sustain a phenomenal holism of the complete and necessary variety, what can? Can anything? The answer, I think, is 'quite possibly', and indeed, we have already encountered a promising-looking route forward. When speculating earlier (in section 3) as to what might have led Sprigge to endorse the complete phenomenal interdependence thesis, I suggested that he may have been influenced by the following line of thought: (1) the sort of unity which we find in consciousness – *phenomenal unity* – is of a very distinctive kind; (2) being unified in this way impacts upon the intrinsic phenomenal character of the experiences concerned; (3) the resulting difference in phenomenal character is also of a very distinctive kind, for the influence that each part derives from the whole reflects the character of the other parts which compose the whole; (4) hence this influence cannot be replicated by other types of whole. I also pointed out that if we suppose, as seems plausible, that all parts of a total conscious state are phenomenally unified, then any unity-generated phenomenal effects will be felt by all the experiential parts of such wholes. The resulting holism will be both complete, and necessary: complete because it applies to all parts of the relevant class of phenomenal wholes, necessary because it is a direct product of phenomenal unity, which is itself an essential (or defining) feature of such

wholes – a collection of experiences forms a phenomenal whole if, and only if, all members of the collection are phenomenally unified.

The general line of argument looks promising, but for it to be viable the holist needs to be able to mount a persuasive case for supposing that collections of individual experiences that are phenomenally unified necessarily acquire distinctive phenomenal features *by virtue of being so unified*. In chapter 9 of *Stream of Consciousness*, I proposed one route to this conclusion – a route which still strikes me as having some merit.

The point of departure is a particular way of thinking of phenomenal unity. Suppose we confine our attention to what can be said about the unity of consciousness at a purely phenomenological level. If I hear a bell ringing while looking at the tree outside my window, these two phenomenal contents – one auditory, one visual – are undeniably *unified*, and they are unified by virtue of being *experienced together*. Let us call the relationship that phenomenal objects (or contents, or properties) have when they are experienced together in this way, 'co-consciousness'. From a purely phenomenological standpoint, co-consciousness looks to be a primitive relationship, at least to the extent of being direct and unmediated. In the case of my seeing the tree while hearing the bell, I experience the visual and auditory contents together, but this 'togetherness' does not seem to involve the experiencing of any additional experiential ingredient which comes between the contents, and serves to connect or bind them. There are simply the auditory and visual contents, together in my consciousness.

As soon as our attention is drawn to the existence of the co-consciousness relationship, it is evident that it is a familiar, if easily overlooked, feature of our conscious lives. Visual contents can be co-conscious with auditory contents, but visual contents can also be co-conscious with olfactory contents, bodily sensations, conscious thoughts, feelings, and so forth; in a similar fashion, any auditory content can be co-conscious with phenomenal contents of these and other kinds. More generally, any two parts of our total conscious states at any given time are co-conscious. More generally still, it is plausible to think that the same applies irrespective of how we opt to divide these total states into parts. So, for example, the contents figuring in the left half of my visual field are co-conscious with the tingle in my toe and my current conscious thoughts; but so too are the contents in the *whole* of my visual field, and the same applies to the contents in just the lower third of the field: these too are also co-conscious with the toe-tingle and my conscious thoughts. Precisely the same applies for cross-modal parts. The experience

which consists of my conscious thoughts and the left half of my visual field is itself co-conscious with the vague ache in my lower back, and the sensations of pressure in the sole of my foot.[20] Co-consciousness is thus a *pervasive* relationship, in this sense: no matter how we choose to divide a total conscious state into parts – irrespective of how complex or simple these parts happen to be – all of these parts are connected to one another by the co-consciousness relationship.

Turning matters around, the co-consciousness relationship provides us with a natural way of defining a complete or total conscious state. Such a state is simply one whose parts are all mutually co-conscious. More precisely, a total conscious state is a collection of experiences (or experiential parts) which are all co-conscious with one another, and which are not parts of any larger collection of experiences whose members are all co-conscious with one another. To put it another way, a total experience is a *maximal* collection of mutually co-conscious experiences. It is plausible to think that our ordinary experiences, at any one time, form parts of total experiences in the sense just defined.[21]

The next question is whether phenomenal contents which are unified in this manner acquire any additional phenomenal characteristics as a result – characteristics which they would not and could not acquire in any other way. To simplify, let us suppose that your total state of consciousness at the present time is confined to the hearing of a bell ringing, and the seeing of a tree. We can label these auditory and visual contents 'F-type' and 'G-type' respectively, and use the expressions 'a_1' and 'v_1' to refer to the token experiences involved.

[20] Those who subscribe to the doctrine of 'unrestricted composition' in the physical realm hold that every combination of material items, no matter how scattered or disparate (from the standpoint of common sense) counts as a fully legitimate *physical object*. In an analogous manner, I count any combination of parts in a total conscious state as 'an experience' – even if many of the resulting experiences are of unfamiliar kinds.

[21] It is very natural to think that our own experience at any given time forms a total state defined in this way. There are those, however, who argue that a single subject at a single time could have three experiences E_1, E_2 and E_3, which are such that E_1 and E_2 are co-conscious, E_2 and E_3 are co-conscious, but E_1 and E_3 are not – cf. M. Lockwood, *Mind, Brain and the Quantum* (Oxford University Press: Oxford, 1989), chapter 6. If co-consciousness is a transitive relationship, as I (tentatively) argue (*Stream of Consciousness, op. cit.*, §4.5 and *The Phenomenal Self* (Oxford: Oxford University Press, 2008), §8.6), then experiential structures of this sort are impossible, and experiences can only partake in wholes whose parts are all mutually co-conscious.

Is the character of a_1 affected in any way by virtue of being co-conscious with v_1? In a real-life case, the purely auditory character of a_1 may well be largely unaffected by virtue of being experienced together with v_1, but bearing in mind the possible prevalence of cross-modal interactions, it would be wrong to insist on this. To simplify still further, let us set to one side the contingent interference effects of the kind we encountered in sections 4 and 5. The stipulation that there are no cross-modal or gestalt-based interference effects – at least in this particular instance – serves the useful purpose of allowing any purely unity-generated phenomenal effects or influences to stand out in clear relief.

In clarifying the situation further it will help to distinguish between two kinds of phenomenal feature. Let us call the purely and exclusively visual features of v_1, and the purely and exclusively auditory features of a_1, the *local* phenomenal properties of these experiences. In addition to these local features – in this case, an F-type auditory content and a G-type visual content – these token experiences each possesses *relational* properties of a phenomenal kind, properties each experience possesses by virtue of being experienced together with other token experiences. We can refer to these unity-generated features as *global* phenomenal properties. Since we are considering a total state of consciousness of an unusually simple kind, the global properties of a_1 and v_1 are also very simple: a_1 has the global property of 'being co-conscious with a G-type visual experience', whereas v_1 has the global property of 'being co-conscious with an F-type auditory experience'. Of course, in the more realistic case of a more complex total state of consciousness, the global properties of a typical part of this total state would be considerably more complex. For the total state *S* composed of experiential parts P_1, P_2, P_3, P_4 … P_N, of local phenomenal types T_1, T_2, T_3, T_4 … T_N respectively, the global character of (say) P_2 would be along the lines of 'is co-conscious with experiences of local types T_1, T_3, T_4 … T_N'. But since this additional complexity does not affect the essentials of the situation, we can safely remain with our simple example.

This stage-setting out of the way, we can move on to the key question: if we want to specify the phenomenal character of a token experience such as a_1, does it suffice to mention just its local phenomenal properties (in this case, F-type), or do we also need to include its global phenomenal properties as well? If the specification is intended to be maximally complete, capturing as much as possible, there are strong grounds for supposing that it will have to encompass both local *and* global properties. The complete story about a_1's phenomenal features will obviously include mention of its local properties

(it is F-type: the hearing of a bell), but if this story is to be truly complete it cannot end there, for it will also have to include the *relational properties*, of a phenomenal sort, that a_1 possesses (*i.e.* 'is co-conscious with a G-type visual experience'). A truly exhaustive specification of the physical properties of any material object will obviously include the object's intrinsic and relational properties – *e.g.* its distances from other objects, the forces acting on it – why should it be otherwise with phenomenal objects?

Adopting this inclusive policy also has a clear phenomenological rationale. A complete and accurate account of the phenomenal character of any experience will exactly capture *what it is like* to have that experience. The nature of phenomenal unity is such that a complete and accurate account of the character of a_1 which failed to include reference to the occurrence of a G-type visual experience would simply be failing *fully* to capture what it is like to experience a_1. The experiencing of a_1 involves a hearing of a ringing bell, but not just that: it involves a hearing of the sound of a ringing bell *that is co-conscious with the seeing of a tree* – this relational element is an important ingredient in what it is like to have this particular auditory experience. It is for this reason that we cannot hope to capture the complete phenomenal character of a_1 without including both the local and the global phenomenal properties that it possesses. And of course, what goes for a_1 goes for any other token experience that is part of a larger experience.

Although it may not be immediately obvious, the distinctive nature of the co-consciousness relationship is playing a crucial role here. As noted above, co-consciousness is not some additional element in experience that possesses phenomenal features of its own. When a_1 and v_1 are experienced as co-conscious, they are simply experienced together: there is no trace of any *experiential glue* (as it were) binding or connecting or linking them. Two contents that are experienced together are in *immediate phenomenal contact* (as we might put it) with one another. This contact is of a distinctively pervasive kind: a_1 and v_1 are not only experienced together, but every part of a_1 (the auditory experiencing of the bell ringing) is co-conscious with every part of v_1 (the visual experiencing of the tree). It is because of the peculiar – but very familiar – intimacy created by the co-consciousness relationship that such a strong case can be made for supposing that a maximally complete and revealing characterization of what it is like to have either of these contents must make mention of the other.

If we do opt to include both local and global properties in our specifications of the phenomenal character of token experiences, the

consequences are far-reaching and dramatic. Since the proper parts of any total state of consciousness are all mutually co-conscious with one another, it follows that a complete specification of the phenomenal character of any one of these parts will make essential reference to all the other parts. As we saw above, in the case of the total state *S* composed of P_1, P_2, P_3, P_4 … P_N, the global character P_2 would be along the lines of 'is co-conscious with experiences of local types T_1, T_3, T_4 … T_N', and similarly *mutatis mutandis* for the other parts. In this manner, the character of the whole impacts on the character of each of its parts – or if you prefer, the character of each (proper) part of a phenomenal whole – impacts upon the character of all the other parts of the same whole. Once we have recognised that the global phenomenal properties of an experience contribute to its overall phenomenal character, then provided we opt to individuate token experiences in the standard way, by holding that the precise phenomenal character of any experience is essential to it, we have the result that it is impossible for any token experience to exist in a total state that is of a different type to the one to which it actually belongs.[22] Given that this form of interdependence applies to the constituents of all total states, irrespective of how they are divided into parts, we have a phenomenal holism that is complete. Since the influence on phenomenal character derives from the co-consciousness relationship – the very relationship which binds experiences into phenomenal wholes in the first place – the holism in question is of the necessary variety: it extends to all parts of all total states of consciousness, in all possible worlds.

I think that a defence of phenomenal holism along these lines, rooted as it is in the distinctive unity that is to be found in our streams of consciousness, would at the very least have been congenial to Sprigge. Indeed, there are grounds for thinking this is close to (part of) what he had in mind. In a passage of the *Vindication* dealing with these matters, he considers an experience that is part of a particular total state of consciousness, and tells us that this experience will lack a precise phenomenal character that is independent of the whole to which it belongs, because imagining the experience 'in its full nature is necessarily imagining it as just that aspect of just such a total experience'. It follows that experiential parts cannot coherently be envisaged as existing in wholes that are of a

[22] To simplify matters I am overlooking here an important distinction between type-specific holism and token-specific holism – for a fuller treatment see *Stream of Consciousness*, *op. cit.*, §9.2 and also *The Phenomenal Self*, *op. cit.*, §9.5.

different type to those wholes to which they actually belong. He continues:

> such components may share quite specific characteristics with components in centres [of experience] not duplicating the centre to which it itself belongs, but the full possibility of which it itself is an actualisation is bound to be different from that actualised by such another component, not because one decides optionally to include its relational characteristics as part of its individual essence, but because its inherent character and its relations to the rest are not separate matters at all.[23]

Sprigge may not have used the *terminology* of 'local' and 'global' properties, but this distinction is certainly implicit in this passage, as is the claim that both sorts of property must enter into fully adequate specifications of the phenomenal characteristics of experiences. And as we have just seen, if this claim is correct, then a surprisingly strong and wide-ranging form of phenomenal holism swiftly follows.

University of Liverpool

[23] T. L. S. Sprigge, *Vindication*, *op. cit.*, 170–1.

Is There a Metaphysics of Consciousness Without a Phenomenology of Consciousness? Some Thoughts Derived from Husserl's Philosophical Phenomenology

EDUARD MARBACH

Abstract
The paper first addresses Husserl's conception of philosophical phenomenology, metaphysics, and the relation between them, in order to explain why, on Husserl's view, there is no metaphysics of consciousness without a phenomenology of consciousness. In doing so, it recalls some of the methodological tenets of Husserl's phenomenology, pointing out that phenomenology is an eidetic or *a priori* science which has first of all to do with mere ideal possibilities of consciousness and its correlates; metaphysics of consciousness, on the other hand, has to do with its reality or actuality, requiring an eidetic foundation in order to become scientifically valuable. Presuming that, if consciousness is to be the subject-matter of a metaphysics which is not simply speculative or based on prejudice, it is crucial to get the phenomenology of consciousness right, the paper then engages in a detailed descriptive-eidetic analysis of mental acts of re-presenting something and tries to argue that their structures, involving components of non-actual experiencing, pose a serious problem for a materialistic or physicalistic metaphysics of consciousness. The paper ends with a brief comment on Husserl's broader view of metaphysics, having to do with the irrationality of the transcendental fact, *i.e.* the constitution of the factual world and the factual life of the mind.

Let me first briefly comment on the perhaps somewhat puzzling title of this paper. The philosophical background of my interest in consciousness stems foremost from Edmund Husserl's *phenomenology*. The term 'phenomenology' (not unlike other terms with the suffix '-logy', *e.g.* 'bio-logy', 'geo-logy', *etc.*) in the present context refers to a theoretical enterprise; it is not used to designate phenomenal aspects of experiences, their so-called 'phenomenal', or 'subjective', or 'qualitative' character, or their ways of seeming or feeling for someone, *etc.*, as is often the case in analytical philosophy of mind and consciousness. Husserlian phenomenology can in a first approximation be seen as a reflection-based descriptive science concerned

doi:10.1017/S1358246110000123

with the *essence* of the *phenomena* of consciousness together with the objects of consciousness as such, *i.e.* as *correlates* of consciousness. Taken in this sense, the *phenomenology* of consciousness might appear to be *all there is* to be philosophically investigated with regard to *consciousness and its objects*. However, Husserl also had a view on the relationship between phenomenology and metaphysics.[1] Accordingly, I propose to address the main question, 'Is there a metaphysics of consciousness without a phenomenology of consciousness?' in two sections, followed by a short coda. In the first section, I explain why the short answer, 'no, there isn't', seems to me to be correct by expanding a bit on the question 'why not?' Doing this requires some comment on Husserl's conception of philosophical phenomenology, metaphysics, and the relation between them; obviously, I cannot provide here more than a sketch of this. In section 2, I outline a difficulty I have with materialistic or physicalistic approaches regarding, as it is often put, 'consciousness and its place in nature' or, put another way, with 'the ontological status of consciousness relative to the world of physical reality', a difficulty that arises in my understanding as soon as I take into consideration detailed analyses of certain conscious experiences, particularly cognitive experiences, along a Husserlian approach to the phenomenology of consciousness, as sketched in section 1. In conclusion, I introduce a twist regarding Husserl's view of metaphysics by very briefly putting forward what he characterized as 'metaphysics in a new sense' which may to some extent at least meet concerns that were at the heart of Timothy Sprigge's metaphysical interests.

1. The Relationship Between Phenomenology and Metaphysics

In order to characterize Husserl's view of the relationship between phenomenology and metaphysics, it may be helpful to recall some of the methodological tenets of his phenomenology. Husserl himself, *e.g.* in his programmatic Inaugural Lecture at the University of Freiburg in 1917, spoke of his enterprise of descriptive 'pure phenomenology' as 'the science of pure consciousness', or as 'science of the pure phenomena', though obviously not taking 'science' in the sense of the empirical natural sciences which rely on objective, third-person data. Rather, he viewed pure phenomenology

[1] In this context, I am also reminded of William Seager's *Metaphysics of Consciousness* (London and New York: Routledge, 1991).

as 'a new philosophical basic science'. Husserl's methodology crucially relies on rigorously confining the analysis to that which *reflection* upon *experiences*[2] *themselves and purely as such* provides – qua *consciousness of something* of one kind or another, exclusively taken as *intentional correlate* of the corresponding conscious experience, with a view to elaborating *concepts* of the very *possibility in principle* of experiencing this or that in such and such a way. Such properly *phenomenological concepts* bring forth certain *a priori constraints* concerning possible explanations of the structures of conscious experiences, and lawful dependencies among them.

In this way, I think, Husserl was able to secure a *pure givenness* of his research domain as a *thematically independent* field of investigation. More technically speaking, with the method of what he termed 'phenomenological reduction', he aimed at distinctly delimiting the research domain of phenomenological analysis in its characteristic ownness, that is, at establishing a theme of investigation unmixed with empirical matters of fact. Thus, he left behind the commonsensical everyday conception of conscious experiences as psychological data ascribable to this or that creature, oneself included, understood as this or that empirical self. Moreover, for his *philosophical theme of investigation*, he also set aside the natural scientific conception of experiences as ultimately neurological processes in the brain. As a consequence of this methodologically motivated restriction, one of the most often recurring expressions in Husserl's writings is that of considering 'consciousness purely as it itself' ('*Bewusstsein rein als es selbst*'), namely just as it can be given in *pure* reflection.

To be sure, in the course of the very formation of phenomenological concepts, a given conscious experience of something provides the experiential basis for the sought-after description of its invariant structure or form according to its very possibility (that is, in Husserlian terms, according to its '*eidos*' or 'essence'). Thus, for example, a conscious experience of imagining a flying elephant, a case of recollecting an episode from one's own life, an experience of picturing something, *etc.* will be submitted to such analysis. Of such experiences we all have an everyday knowledge of acquaintance that is reflected in the mental vocabulary of ordinary languages. In a way, then, as Husserl occasionally says, we all 'know' of essential differences of being conscious; however, this knowledge is only

[2] In Husserlian phenomenology, the term 'experience', 'Erlebnis' or 'Bewusstseinserlebnis', covers sensory as well as cognitive, emotional, affective experiences.

implicit, and it is just the task of phenomenological reflection and eidetic analysis systematically to explicate distinct phenomenological concepts of the various kinds of consciousness. Based on one's everyday familiarity with a conscious experience of a certain kind as designated in ordinary language, a factually chosen case will be taken *as a purely arbitrary example* of its kind, a mere starting point for the analysis. Regarding this methodological step, Husserl liked to refer to the mathematicians' way of starting their analyses by saying, 'there are … ' ('*es gibt* …'), say, such and such geometrical figures, prime numbers, *etc*.[3] Similarly, Husserl suggested, the phenomenologist adopts the attitude of saying, 'there is, say, an experience of imagining something', of 'picturing something', *etc*. The chosen experience, forming in this sense nothing more than an arbitrarily selected example, does not bind the phenomenologist *qua* this or that particular subjective experience, existing as a psychological matter of fact which is such and so determined, occurring for example with this or that degree of vivacity and distinctness of content, *etc*. The irrelevance of the psychological matter of fact *as such* for the purpose of the phenomenological concept formation proper can also be seen when we realize that we must engage in a process of varying the conditions in order to define which ones are invariably required, or essential, for making the experience *possible* as against those that can be changed *without* altering the essential structure of the experience *qua* experience of the kind now to be reflectively differentiated from other kinds.

Phenomenological analysis, then, is only interested in truly constituent parts or properties capable of being distinguished in reflection as belonging to, and making together up, a unity of the conscious experience under study in its own essence or nature, *i.e*. in accordance with the conditions of the possibility of its occurrence, and not of the actuality in its variability as a psychological matter of fact. As Husserl put it in a lecture course of 1907:

> The conditions of the 'possibility of experience' are the first. Conditions of the possibility of experience signify, and may signify, here, however, nothing else than all that resides immanently in the essence of experience, in its *essentia*, and thereby belongs to it irrevocably. The *essence* of experience, which is

[3] See, *e.g*. E. Husserl, *Phenomenology and the Foundations of the Sciences*, translated by T. E. Klein, Jr. and W. E. Phol (The Hague: Nijhoff, 1980), §8, 41. See also E. Husserl, *Erfahrung und Urteil. Untersuchungen zur Genealogie der Logik*, edited by Ludwig Landgrebe (Hamburg: Meiner, 1985), §96.

> what is investigated in the phenomenological analysis of experience, is the same as the *possibility of experience*, and everything established about the essence, about the possibility of experience, is *eo ipso* a condition of the possibility of experience.[4]

Clearly, Husserl understood phenomenology as an *eidetic science*, or an *a priori* science of pure consciousness and its correlates, in contrast to the sciences of matters of fact. However, he did not simply limit phenomenological philosophy to the *essence* or the *a priori*. Whereas, in his view, *eidetic phenomenology* has to do with *mere ideal possibilities*, it is *metaphysics* which has to do with *reality or actuality* ('*Wirklichkeit*'). Husserl characterized 'metaphysics' as 'the genuine ('*eigentliche*') science of reality' ('*Realität*') and he made it clear that it was, in the end, also his aim to elaborate a metaphysics, but he would stress that he wanted it to be 'in a serious sense a scientific' metaphysics, as opposed to mere metaphysical speculation, and that therefore he put all his effort *first* into the '*eidetic foundation*' of such a metaphysics.[5]

Husserl had *two reasons* for the view that pure or transcendental phenomenology qua *eidetics* had to *precede*, as its scientific basis, '*metaphysics*' considered as the 'absolute science of the factual reality' ('*faktische Wirklichkeit*'). *On the one hand* – and I think that this is a truly important point for all studies of consciousness -, consciousness has a very peculiar 'nature' or essence: consciousness, conscious experience, is in an 'incessant flux', in an (as Husserl often put it) 'eternal Heraclitean flux'. As pure consciousness, it is, therefore, scientifically graspable on the basis of *essences* only. Merely living through individual experiences from moment to moment and reflecting upon them, in this sense having *experience* ('*Erfahrung*') of consciousness is not enough for a phenomenological *science* of consciousness which must provide results that are intersubjectively controllable. To put it succinctly, 'the fact' of this or that occurring conscious experience which serves as the basis for the reflective investigation, is here – in pure reflective phenomenology – 'determinable *as that of its essence* and only *through* ('*durch*') *its essence* and in no

[4] E. Husserl, *Thing and Space. Lectures of 1907*, translated by Richard Rojcewicz (Dordrecht: Kluwer, 1997), §40, 119.

[5] Compare draft of a letter to Karl Joël, 11. III. 1914 in E. Husserl, *Briefwechsel, Band VI. Philosophenbriefe*. In Verbindung mit Elisabeth Schuhmann, herausgegeben von Karl Schuhmann (Dordrecht: Kluwer, 1994), 205f.

way to be documented by means of inductive experience as it is the case in the objective world'.[6]

On the other hand, the reason for the view that phenomenology as eidetics (as science of the possibilities) had to precede 'metaphysics' was for Husserl linked to a general principle of the theory of science, namely that everything factual ('*alles Tatsächliche*'; all matters of fact, all *a posteriori*) finds *ultimately* its full rationality in the eidetic or in the *a priori*. As Husserl held – very much in line with the traditional so-called rationalists and their 'vérités de raison' as against 'vérités de fait' (Descartes, Leibniz) – all reason (rationality) in the *a posteriori* has its principles *a priori*, and these principles are the justifying reasons ('*Rechtsgründe*') of the objective validity of knowledge. The old ontological doctrine, that knowledge of the possibilities has to precede knowledge of the actualities ('*Wirklichkeiten*') is, according to Husserl, 'a great truth' if correctly understood and put to use in the right way.[7]

In this sense, then, Husserl considered eidetic phenomenology of the possibilities to be *First Philosophy*, and metaphysics, as the science of the actualities, is 'Second Philosophy' or 'Empirical Philosophy'. By way of applying pure eidetic phenomenology to matters of fact concerning consciousness – in analogy to applying pure mathematics to the actual world of physics –, empirical research into consciousness would receive its ultimate, *i.e.* its *metaphysical* interpretation, behind which it would make no sense to look for a further interpretation.[8] In short, then, from the point of view of Husserl's philosophical phenomenology there can't be a scientifically valuable *metaphysics* of consciousness without a *phenomenology* of consciousness.

[6] Compare E. Husserl, *The Crisis of European Sciences and Transcendental Phenomenology*, translated by David Carr (Evanston: Northwestern University Press, 1970). §52, 178; translation slightly amended.

[7] See, *e.g.* E. Husserl, *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Erstes Buch. Allgemeine Einführung in die reine Phänomenologie*, edited by Karl Schuhmann, *Husserliana* III/1 (Den Haag: Martinus Nijhoff, 1976), 178; compare also *Einleitung in die Logik und Erkenntnistheorie. Vorlesungen 1906/07*, edited by U. Melle, *Husserliana* XXIV (Dordrecht: Martinus Nijhoff, 1984), §40, 236ff.

[8] See E. Husserl, *Erste Philosophie (1923/24). Kritische Ideengeschichte*, edited by Rudolf Boehm, *Husserliana* VII (Den Haag: Martinus Nijhoff, 1956), 187f., note 1.

2. A Difficulty Regarding Consciousness and Its Place in Nature

Let us now turn to the above-mentioned difficulty regarding the place of consciousness in the natural world, or put another way, regarding the ontological or metaphysical status of consciousness relative to the world of physical reality. In order to clarify this difficulty, it is necessary to engage in some descriptive-eidetic phenomenology.

Despite all my respect and sympathy for defenders of a unitary view of reality inspired by modern science, I continue to be deeply impressed by results from philosophical phenomenology, providing us with forms or structures of consciousness that *in principle* seem to have *no analogue in the natural world* as it is studied by the sciences.[9] Indeed, as I understand the lesson of the sciences of the external physical world out there and the neurosciences of the brain, the world as it really is in itself exists as a whole of spatial and/or temporal parts-*outside*-parts, three-dimensional time-slices, all the way down to the ultimate particles on a scale of nanometres, forming cohesions and functional unities involved in all kinds of states that are parts of a system of possible states. By contrast, as Husserl often observed, it belongs to the essence of conscious life *not* to be structured in terms of 'the spatial outside one another, into one another and through one another and spatial totality' but to contain in itself intentional implications and modifications that cannot be accounted for in spatial terms, even though spatial metaphors in describing forms or structures of conscious experiences crop up time and again.[10] In Husserl's view, even Brentano was still a naturalist precisely because 'something like *intentional implication* and intentional analysis as analysis of possibly continuously intertwined sense-giving he did not yet know'.[11]

Clear examples providing evidence for such structures *sui generis* are given by certain conscious experiences that involve *non-actual components* or *moments* in their very way of establishing intentional reference to something when they *actually* occur. Such a finding, I

[9] See E. Husserl, *Phänomenologische Psychologie. Vorlesungen Sommersemester 1925*, edited by Walter Biemel, *Husserliana* IX (Den Haag: Martinus Nijhoff, 1962), 35ff.

[10] See, *e.g. ibid.*, 37: '… dass es zum Wesen des Bewusstseinslebens gehört, anstelle des räumlichen Aussereinander, Ineinander und Durcheinander und räumlicher Ganzheit ein intentionales … ineinander meinend Beschlossensein in sich zu bergen…'.

[11] *Ibid.*

take it, is very hard to integrate into a description of reality couched in the language of physics or the neurosciences which a complete naturalization of consciousness would have to amount to. The conscious experiences in question make up the large and important class of mentally re-presenting something that is not actually present, '*Akte der Vergegenwärtigung*', acts of re-presenting something, which Husserl has investigated extensively. Thus, *e.g.* when I am engaged in an activity of visually re-presenting, thereby actually intentionally referring to something – say, a past event that I remember visually, or a possible one that I imagine, or an object that I see as being depicted in a picture – then not only is the *re-presented object or event* consciously given as being *not actually present*, but at the same time the conscious experience of so referring to something absent contains *within its very structure an experiential component of not actually performed perceiving (seeing)*. If so, then there is something actually occurring in this world an essential property of which is that parts of it are consciously experienced *as non-actual*.

By the way, I take it that as long as we would confine our attention somewhat abstractly just to *sensory* conscious experiences occurring here and now in causal interaction with the actual surrounding world, there is not so much reflective work for phenomenology to do regarding structures of these experiences qua structures of consciousness to be articulated in the language of phenomenology. Still, their temporal form, their associative connections and emotional colourings as well as their being part of a *horizonal consciousness* ('*Horizontbewusstsein*') could be, and have been, studied phenomenologically as making up their specific what-it-is-likeness *qua* sensory experiences. Much more structural knowledge would no doubt come from neurophysiological investigation into the specific physical character of the brain states that are causally responsible for the phenomenal character of the corresponding sensory experiences – even though, I think, the so-called 'hard problem' would remain, the problem of explaining why this or that sensory experience should have the specific what-it-is-like aspect it does, rather than having some other phenomenal character or having none at all.[12] I am not so thrilled by this problem at the level of *sensory* experiences given their facticity and hence intractability. It is very tempting to say that I am pleased to have all this motley of sensory impressions, they make life so colorful, but there they are, matters of fact, and,

[12] Cf. *e.g.* D. Chalmers, *The Conscious Mind* (Oxford: Oxford University Press, 1996).

as Tim Crane disarmingly put it, 'some facts are amazing, and sometimes that's all that can be said'.[13]

Let us stick, then, to kinds of more or less *complex non-sensory* or properly *cognitive* acts of re-presenting something, involving also sensory components, to be sure. All such acts, I also take to be in themselves so many modes, simply more complex ones, of what-it-is-like to experience them, to live through them and, by having them, to establish intentional reference to something that we experience in its absence. Concerning the conscious experiences themselves, they are exemplified solely in the experiential life of a subject, originally in my own mental life, and this already when they are unreflected on at the time we have them. This, I take it, is just the sense of the present-day phrase that there is 'something-it-is-like' to have this or that conscious experience: it *pre-reflectively*, as Husserl and later on Sartre would have put it, feels somehow or other, say, to perceive something out there, and it feels differently or there is something-other-it-is-like, say, to see something in a picture, or to remember something, *etc*. Only on the basis of the subject's reflection upon the experiential life can these ways of what-it-is-likeness then be described in general statements regarding their distinct structures, forming thereby descriptive phenomenological concepts of the conscious experiences. To account philosophically for similarities and distinctions among the varieties of the intentionality of conscious experiences and their intentional correlates, as they present themselves as phenomena, *i.e*. at the level of their phenomenality, was no doubt much at the center of Husserl's eidetic 'science of consciousness'.

Let me expand a little on some distinctions with a more refined analysis along Husserlian lines. The case of *picture viewing* is particularly well suited to point out what differently structured conscious experiences as such are able to achieve. Consider, for example, a change from a conscious experience of simply perceiving a green pine over there in the meadow to a conscious experience of *pictorially re-presenting* a green pine and its surroundings. Here, reflective analysis, and I think, at first at least, *only* reflective analysis, is able to articulate a radical difference regarding the structure of the corresponding conscious experiences themselves and, correlatively, of the ways of givenness of the intentional object, the green pine and its surroundings. To begin with, notice the following aspect. Concretely viewed, *i.e*. not in isolation, the color green in the pictorial

[13] T. Crane, *Elements of Mind. An Introduction to the Philosophy of Mind* (Oxford: Oxford University Press, 2001), 89.

representation is attributed, *not* to a portion of the flat canvas as such, but rather either to the merely pictorially appearing but not actually given pine on the canvas or to the depicted and again not actually given pine somewhere in the real world or in a fictional space, depending on where the person viewing the picture is focusing her attention. Either way, the presently manifest quality green is here *taken*, or *apprehended*, *as a property of an object that is not actually present* over there where the canvas actually is and where the phenomenal quality green appears. If so – namely in so far as I have conscious experience of the green as of a property of the pine – the phenomenal quality green cannot be given solely in virtue of a sensory-perceptual visual experience; for there is actually no pine to be visually perceived over there on the wall where the painting is hanging. What we have instead is a more complex conscious experience of intentionally referring to the depicted pine by way of seeing as it were the green pine in so far as it appears on a canvas, *i.e.* as a pictorial object. In its unified structure such an experience also involves as a conscious component among others my *seeing* the green pine *in the mode of non-actuality*, namely precisely in so far as the *pine* seems to appear on the canvas over there that I actually see. In other words, the phenomenal quality green appears no longer as actually belonging to a portion of the canvas or the picture-surface over there. But the pictorially appearing pine's being green – in contradistinction to a perceptually appearing pine's being green – only appears so because of my consciously taking the pictorial object 'pine' to be a representation of an absent (real or imaginary) tree that is only depicted in the picture that I actually see in the present surroundings.

Crucially, the conscious component of my *seeing in the mode of non-actuality*, as Husserl has convincingly explained, *constitutes an intentional modification* and can be said to be *intentionally implied* in the unified conscious experience that I have of seeing something pictorially represented, and clearly this is not a matter of being *spatially* contained within the experience in the sense of being a part outside other parts. Let me spell this out in a little more detail. What other components are there in such a unified experience of pictorially representing something that do not stand in external spatial relations to one another? Consider first that while I am intentionally referring to the green pine only in so far as it appears in the picture over there, I am simultaneously aware of my present surroundings here and now in virtue of actually perceiving the world around me as well, although somewhat less attentively as long as I am interested in the depicted scene or in the scene as it pictorially appears on the canvas. The fact that I take the pine and its surroundings not to be actually

present over there depends, precisely, on a consciously experienced contrast concerning, on the one hand, the ways of givenness of the objects around me, and among them the picture on the wall showing the green pine in its surroundings, and on the other hand, the ways of givenness of the pine either in so far as it pictorially appears on the canvas or is taken as the depicted pine somewhere in the real world or in some fictional world, respectively. There is a *clash* between that which I take to be actually real around me and that which I take to be just a pictorially represented, not actually given, scene. It is in relation to this pictorially appearing or depicted scene in contrast to the actually perceived surroundings that I speak of seeing something in the mode of non-actuality; for I cannot actually see what is not actually there. To be sure, I actually look over there to the picture and thus, in a sense, I *actually* see something over there, just as I actually see the wall over there where the whole picture that I actually see is hanging – I actually see patches of green *etc*. However, as Husserl would urge, as soon as I take those green patches to be ways of appearing of a pine – be it the pine in so far as it appears in the picture or the depicted real or imaginary pine – the very experience of seeing is consciously modified: I do not believe I actually see a pine over there. As I understand the matter, my seeing the pine is, however, not simply a particular kind of seeing, a so-called 'representational seeing' or a 'seeing-in' as 'the subsidiary capacity that we also have', besides 'the standard capacity that we have of vision'.[14] Instead, to the extent that in picture viewing vision is involved in relation to the pictorially appearing objects, it is altogether experientially modified, not unlike what obtains in visually remembering or imagining something that is not actually, presently given, although not quite like it either.

To clarify matters further, notice next that with my seeing as it were the green pine in its surroundings in so far as it appears in the picture or is depicted by the picture that I actually see over there, I do not simply intentionally refer to the pine as if I were seeing the pine *itself* in some imaginary or remembered place. Yet, there is similarity to some extent between picture viewing and imagining or remembering, namely to the extent, precisely, of involving a visual experience in the mode of non-actuality and an object that is not actually given. For notice that in the case of, say, an experience of simply

[14] See, *e.g*. R. Wollheim, *The Mind and Its Depths* (Cambridge, Mass.: Harvard University Press, 1993), 188 and C. Peacocke, 'Depiction', *The Philosophical Review* **96:3** (1987), 383–410, discussing earlier proposals by Wollheim.

imagining a pine in imaginary surroundings, the pine would again not actually be given but only appear *as it were*, this time in relation to an imaginarily occupied point of view of mine from where I would see the tree *itself*, albeit only as it were, *i.e.* in a consciously modified manner. This is to say that once again vision would not be functioning in a standard way. My visual experience would indeed not actually be performed but only be re-presented: it would be as if I were seeing the not actually given pine from some place, at some distance and orientation relative to my standpoint in the imaginary space. By contrast, when I intentionally refer to the pine in so far as it appears in the picture, I can do so only because at the very same time I am aware of the pine's merely being a pictorial representation *of* a tree, be it a real tree in the world or be it merely a fictional tree. The situation, in picture viewing, is indeed rather special, as has often been observed by philosophers and also among scientific psychologists, among whom may be numbered Richard Wollheim, Richard L. Gregory, J. J. Gibson and others. I am neither simply actually perceiving or seeing an object 'in' a picture, nor am I simply purely mentally representing the not actually given object itself, say, in remembering or imagining it in a different space and at a different time. Rather, to sum up, in picture viewing, *i.e.* when I pictorially re-present some absent object – a thing, person, place, or event – I am intentionally referring to the absent, not actually given object (the so-called *sujet*, say, of a painting) by means of my re-presenting a non-actual perceiving of the object (the *sujet*) in so far as it appears in the actually perceived picture while simultaneously presenting my actual surroundings here and now.

In such conscious experiences of purely mentally or of pictorially re-presenting something absent (past, future, merely possible, or elsewhere in space, *etc.*), as well as in more complex conscious experiences of mentally re-presenting something absent by way of re-presenting an act of pictorially re-presenting it (say, in remembering or imagining an episode of picture viewing), there is always the reflective phenomenological finding of those modified components of *non-actuality* belonging to the very structure of the actually experienced mental activity: one's being conscious or aware of *intending* something absent, and one's being *innerly* conscious or aware of the non-actual experience – in the simplest case, a perceptual experience – that is implied in the actually occurring establishment of the intentional reference to the absent object. The very givenness of the intended absent, not actually given object as well as the very givenness of the intentionally implied or modified non-actual component of, say, seeing as it were the absent object, is bound to occur within the unified

structure of the actually performed conscious experience of re-presenting together with the presentation of the actual surroundings.

In this sense, then, I think that with each instance of an actually occurring conscious experience of re-presenting something absent in virtue of components that are *experienced as non-actual*, we have clear examples of actually existing entities in our world that do not fit into the framework of the sciences of the natural, physical world. At any rate, those non-actual components of the actually occurring experiences of re-presenting something *cannot* be understood as being *spatially distinct* from the actual experiences; being non-actual, they can only be individuated by the experiencing subject as being consciously modified re-presented parts of the subject's actual re-presenting something.

3. Coda: Metaphysics in a New Sense

In Section 1, the relationship between phenomenology and metaphysics was presented such that *metaphysics* would only be at stake in Husserl's overall conception of philosophy as 'second philosophy', or 'empirical philosophy' in the sense of an *application* of *eidetic first philosophy* to empirical matters of fact in view of an ultimate interpretation, or a rationalization, of the *a posteriori*. In conclusion, I would like to point out, albeit very briefly, that Husserl's philosophy of the actuality ('*Wirklichkeit*'), *i.e.* his '*metaphysics*', was *ultimately* not simply equivalent to such an eidetic phenomenological interpretation of the sciences of empirical matters of fact. What Husserl, from relatively early on in his work, did consider beyond has to do with '*the irrationality of the transcendental fact* which expresses itself in the constitution of the factual world and in the factual life of the mind: thus *metaphysics in a new sense*'.[15] In Husserl's view, the *irrational fact of the rationality* of the world was the object of *metaphysics* in this new sense.

Intriguingly, in his last years, Husserl even questioned his distinction between First and Second Philosophy, formerly oriented after the priority of the eidetically possible over the actual, eidetic philosophy being considered as science of *pure possibilities*. Even 'the pure I' ('*das reine Ich*'), belonging to the 'field' of transcendental phenomenology, was for many years conceived by Husserl as a pure possibility, an 'I in general' ('*Ich überhaupt*'), preceding the actuality. However, in a text from 1931, he totally reversed this conception by noting, with

[15] See E. Husserl, *Erste Philosophie*, *op. cit.*, Chapter 3: 'Rationalismus und Metaphysik der Neuzeit', 188n.

regard to the *eidos transcendental I*, that 'we have here a peculiar and unique case, namely concerning the relationship between fact and eidos': whereas, generally, the being of eidetic possibilities – *e.g.* in pure mathematics – is free, independent of, the being or not-being of any *corresponding realization* of such possibilities, '*the eidos transcendental I is inconceivable without a transcendental I as factual*'. After speaking of the full ontology as teleology and of the facticity presupposed by this teleology, the text continues: 'We end up with ultimate "matters of fact" – *Urtatsachen*, with ultimate necessities, the *Urnotwendigkeiten*. But *I* think them, I question back … I am the primal fact ('*Urfaktum*') in this course, I see that these and these primal constituents of my own result in my questioning, as primal structures ('*Urstrukturen*') of my facticity. And that I in myself carry a core of "primal contingency" ('*Kern von "Urzufälligem"*') in eidetic forms, in forms of potential functioning, in which the mundane eidetic necessities are then founded. I cannot transgress my factual being and in it not the intentionally contained co-being ('*Mitsein*') of others *etc*. thus the absolute actuality ('*die absolute Wirklichkeit*')".[16]

Even though, after this reversal, Husserl did no more systematically elaborate the relationship between the phenomenological eidetics of consciousness, the actuality and the empirical fact, he firmly upheld the eidetics as condition of the possibility of transcendental cognition of consciousness and its objects until as late as in the *Crisis* of 1936[17]. And even though for us the acknowledgment of a thorough facticity of the transcendental I may raise novel and pressing questions regarding the naturalization of consciousness in a phenomenological perspective, it would still seem that the very *theme* of consciousness as such and of its objects considered as 'correlates' – as it has been the central theme of Husserl's own philosophical phenomenology – remain a valuable theme of philosophical investigation into consciousness, whatever the core of 'primal contingency' in eidetic forms will turn out to be.[18]

University of Bern, Switzerland

[16] E. Husserl, *Zur Phänomenologie der Intersubjektivität, Dritter Teil 1929–1935*, edited by Iso Kern, *Husserliana* XV (Den Haag: Martinus Nijhoff, 1973), Nr. 22, 385f.

[17] E. Husserl, *The Crisis*, *op. cit.*, see note 6 above.

[18] I would like to thank Pauline Phemister for carefully editing and linguistically improving my text; my thanks also go to Pierfrancesco Basile for his help, and to all the organizers for extending the invitation to participate in the Conference in Honour of the late Timothy Sprigge.

Making Sense of Phenomenal Unity: An Intentionalist Account of Temporal Experience

JULIAN KIVERSTEIN

Abstract
Our perceptual experiences stretch across time to present us with movement, persistence and change. How is this possible given that perceptual experiences take place in the present that has no duration? In this paper I argue that this problem is one and the same as the problem of accounting for how our experiences occurring at different times can be phenomenally unified over time so that events occurring at different times can be experienced together. Any adequate account of temporal experience must also account for phenomenal unity. I look to Edmund Husserl's writings on time consciousness for such an account.

1. Introduction

The objects and events we experience have temporal properties such as persistence, change, succession and movement. We can see a flock of birds take flight; feel a persisting, dull, throbbing headache; hear the hum of the traffic outside of our window; feel the softness of the silken fabric as we run our fingers across it. Each of these experiences represents an event with temporal properties. Yet some philosophers have found themselves driven to the conclusion that we never, strictly speaking, experience temporal properties. Thomas Reid claimed for instance that: 'The senses give us information of things only as they exist in the present moment; and this information, if it were not preserved by memory, would vanish instantly, and leave us ignorant as if it had never been'.[1] To experience temporal properties like change, persistence and succession takes time, since change, persistence and succession occur over time. Yet Reid takes our experience to be confined to a 'present moment' that has no duration. Hence he finds himself forced to conclude that we can at best remember what is in the past: 'it is only by the aid of memory that we discern motion, or any succession whatsoever'.[2]

[1] T. Reid, *Essays on the Intellectual Powers of Man*, edited by J. Walker (Boston, MA: Derby, 1855), 211.
[2] *Ibid.*, 237.

doi:10.1017/S1358246110000081

Reid was concerned with our ordinary ways of talking when we make claims like we can see a body move. He takes ordinary talk of this kind to be mistaken; strictly speaking all we can see is the 'present place of an object'.[3] When we see a body move – a horse galloping across the finishing line, for instance – this is because of the contribution of memory. What the ordinary folk are mistaken about is whether they can really see bodies move as opposed to seeing them at particular spatiotemporal locations and recalling where they've just been.[4]

H. A. Prichard also finds error in our folk understanding of auditory perception.[5] He points out that all sounds have duration, and auditory experiences are experiences of sounds. These two claims imply that auditory experience must take up time. Yet Prichard also attributes to the folk the view that individual experiences take place in the present at a single, durationless instant in time. This seems to be inconsistent with the claim that we can experience sounds, but without any experience of sound there can of course be no auditory experience. Nor does an appeal to memory seem to help us in this case. Memory was supposed to supplement experiences confined to a point in time so as to allow us to experience events that seem to extend through time. Prichard's argument

[3] *Ibid*.

[4] Barry Dainton has offered a reply to Reid that is part phenomenological and part empirical ('Sensing Change', *Philosophical Issues* **181** (2008), 362–84.) The phenomenological reply takes as its starting point the seeming truism that we often *see* things moving or changing. This however is what Reid means to deny, so although I don't dispute the phenomenological datum, I can't see why it should dissuade us from agreeing with Reid. The second, empirical response appeals to findings from neuroscience that cells in visual area V5 are dedicated to coding visual information about motion. Lesions to this area can result in a subject no longer seeing smooth, continuous movement. (See J. Zihl, D. von Cameron and N. Mai, 'Selective Disturbance of Movement Vision after Bilateral Brain Damage', *Brain* **106** (1983), 313–40). These findings, while striking, cannot be decisive since it is possible (indeed highly likely) that V5 functions as part of a wider network, and this network may well include areas dedicated to memory of the kind that Reid appeals to. We wouldn't want to say that V5 cut out of a larger brain and placed in a petri dish could realise a visual experience as of motion. Once we allow that motion is coded for by V5 as part of a larger distributed network of activation it seems to me that Reid's theory remains a live option.

[5] H. A. Prichard, 'The Apprehension of Time' in *Knowledge and Perception: Essays and Lectures* (Oxford: Clarendon Press, 1950), 47–51.

seems to establish that we can't make sense of auditory experiences that are confined to durationless points in time.[6]

More recently, Robin Le Poidevin has defended the claim that we cannot be said to experience temporal properties without a contribution from memory.[7] When I hear a note C followed by a note E, for instance, I have an experience of succession. Le Poidevin argues that this is because when I hear E, I also have a 'very recent memory of C'.[8] Thus he seems to agree with Reid that sensory experience must take place in the 'present' where the latter is conceived of as a point in time. Once we accept this claim it seems we have no alternative but to appeal to memory to explain how we can experience events that unfold over time. However if Prichard's reasoning is correct, no such appeal to memory can be made to work in the case of auditory experience.

This paper will attempt to vindicate the folk. We are not guilty of confusion when we say we can see movement, hear sounds, and feel enduring sensations because experiences aren't punctuate, confined to durationless instants. We will see however that the idea of temporally extended experiences can be developed in multiple ways. My aim in this paper will be to argue for an intentionalist account that is Husserlian in provenance. My argument will be that all other theories fail to adequately account for the difference between experiencing properties like succession, persistence and change and undergoing experiences that succeed one another, persist or change. When we experience temporal properties, a relation of phenomenal unity holds that binds together our experiences over time. I shall argue that all accounts of temporal experience – with the exception of the intentionalist one – fail to tell us anything informative about this relation of phenomenal unity. However before I turn to the main argument I must say something in defence of the claim that experiences represent temporal properties.[9]

[6] Ian Phillips makes this point in a brief discussion of Prichard's argument. See I. Phillips, 'The Temporal Structure of Experience' in D. Lloyd and V. Arstila (eds), *Subjective Time: the Philosophy, Psychology, and Neuroscience of Temporality* (Cambridge, MA: MIT Press, forthcoming).

[7] R. Le Poidevin, *The Images of Time: an Essay on Temporal Representation* (Oxford: Oxford University Press, 2007), ch.5.

[8] *Ibid*., 91.

[9] I'm going to assume without argument that experiences have representational content, though I recognise this is controversial in some circles. The issue I want to focus on is whether temporal properties could enter into the representational contents of *experience*, or whether we

2. Do Experiences Represent Temporal Properties?

Things seem a certain way to us in experience because of the phenomenal properties of our experiences. According to the version of intentionalism I favour phenomenal properties are identical with *impure* representational properties, where an experience has an impure representational property by representing an intentional content in a certain manner.[10] My aim in this section will be to show that the events we experience do indeed seem to possess temporal properties because temporal properties enter into the intentional content of experiences. This will still leave us with the question of whether they do so because of the contents of our experiences, or because of our experience's manner of representation. This is a question I return to in the final two sections of my paper.

Susanna Siegel has recently devised a useful method for settling questions about which properties do and which properties don't enter into the contents of experience:[11]

> The method of phenomenal contrast is a way to test a target hypothesis about the contents of experience. The main idea behind the method is to find something that the target hypothesis purports to explain, and see whether it provides the best explanation of the phenomenon… Since contents are supposed to be phenomenally adequate, any target hypothesis will predict that any two experiences differing with respect to the hypothesised contents will differ phenomenally as well. It is thus possible to use this phenomenal contrast as the thing to be explained. The 'target explanation' will say the experiences contrast

---

should agree with Reid when he claims that we can only experience temporal properties through the contribution of some kind of memory.

[10] I owe the terminology of 'impure representational properties' to D. Chalmers, 'The Representational Character of Experience' in B. Leiter (ed.), *The Future for Philosophy* (Oxford: Oxford University Press, 2004). Chalmers contrasts 'impure' representational properties with 'pure' representational properties, properties of having an intentional content. T. Crane, 'Intentionalism' in B. McLaughlin, A. Beckermann and S. Walter (eds), *The Oxford Handbook to the Philosophy of Mind* (Oxford: Oxford University Press, 2009) also defends this distinction.

[11] S. Siegel, 'How Can We Discover the Contents of Experience?', *Southern Journal of Philosophy* **45** (2006), 127–42. Manuscript downloaded from: http://dash.harvard.edu/bitstream/handle/1/3164345/Siegel_DiscoverContents.pdf?sequence=4

> phenomenally because one of them has the hypothesised contents, while the other does not.[12]

The hypothesis we are currently assessing claims that temporal properties like change, persistence, succession and duration can enter into the contents of experience. To apply the method of phenomenal contrast to this case we must find an example of an experience as of a temporal property *p*, and an example in which the same property *p* doesn't figure in experience. This will generate the phenomenal contrast we are interested in. We must then consider whether the hypothesis that experiences represent temporal properties is the best explanation of this contrast. It will prove to be the best explanation if it satisfies the following three features:

> First, the target experience will differ in its sensory phenomenal character from the contrasting experience, rather than differing with respect to a non-sensory category of phenomenal character. Second, the target experience will have representational content. Third, the target experience will have a representational feature – namely, representing the property F, that the contrasting experience lacks.[13]

Consider as a case of phenomenal contrast, C. D. Broad's example of seeing a second hand move around the face of a watch as compared with seeing that the hour hand has moved. In the first case we see the hand moving. In the second case we infer that the hand has moved, perhaps by comparing its position at different times and forming the perceptual belief that it has moved. This example seems to satisfy all three of the conditions we have just stated. There is a difference in what it is like to see the second hand moving and seeing that the hour hand has moved. Only in the first case do we see the smooth and continuous motion of the hand around the watch face. The hour hand moves too slowly for us to see it moving. The two experiences clearly have representational content: the first is accurate if the second hand really is moving, while the second is accurate if we look at the clock twice with an interlude of an hour or more. Finally both experiences represent movement, the temporal property in question.

I will assume that Siegel's method of phenomenal contrast does indeed give us a way of determining which properties enter into the

[12] *Ibid.*, ms, 10.
[13] *Ibid.*, ms, 13.

contents of experience. On the basis of Broad's example we can conclude that temporal properties can enter into the contents of experience. Still, someone persuaded by Reid's reasoning might argue that our experiences can represent temporal properties, only because our experience of what is present to us at an instant in time is accompanied by a memory of the recent past. Obviously not just any kind of memory is going to do the job of securing an experience of succession, persistence and change. In order to contribute to my experience in the right way, memory must represent only events in the recent past. Just how far back in time the relevant memory reaches will depend on just how far back in time my experience seems to reach. Moreover the memory we are appealing to must represent events in the right order. If I experience a series of events A, B and C, at the time when I experience C my memory must also represent B and A in the order they happened. Finally the type of memory we invoke must be passive and involuntary in the same way as experience is passive and involuntary. This differentiates the type of memory involved in the experience of temporal properties from episodic memory, which is subject to voluntary control.[14] Assuming we can give an account of the memory that fulfils all these conditions,[15] we would be on our way to explaining how experiences restricted to a durationless instant could represent events that take time to unfold. A durationless experience could represent temporally extended events by being combined with a kind of sensory memory that possesses the features just enumerated. I'll call this the *memory account*. In the next section I'll argue that even supposing a memory account can be made to work it may still face a major difficulty in accounting for the difference between an experience of succession and a succession of experiences. The resolution of this difficulty will ultimately require us to reject the assumption that experiences are confined to durationless instants.

[14] The desiderata I'm describing here are based on objections Dainton in his *Stream of Consciousness: Unity and Continuity in Conscious Experience*, 2nd edition (London: Routledge, 2006), §5.4 has made to accounts of temporal experience like Reid's that make appeal to memory. I'm also indebted to the discussion in I. Phillips, 'The Temporal Structure of Experience', *op. cit.*, §§ 5 and 6.

[15] For one promising recent attempt see I. Phillips, 'Perceiving Temporal Properties', *European Journal of Philosophy* (forthcoming), §6.

3. Experiencing Succession

Many a discussion of temporal experience has begun with James' observation that 'a succession of feelings, in and of itself, is not a feeling of succession'.[16] To illustrate what James had in mind let us continue with the example of seeing movement. Some movements happen too fast for us to perceive them, as is the case with Locke's example of the 'cannon-bullet' passing through a room taking with it 'any limb or fleshy part of a man'.[17] Other movements happen too slowly for us to see them, as is the case with Broad's example of the movement of the hour hand on a watch. Many movements of objects, however, happen neither too slowly nor too quickly for us to see them. In these cases we see one and the same object, say the second hand on a watch, occupy different positions at different times. Having an experience of the position the second hand occupies at each point in time is having a succession of experiences. This isn't the same thing as experiencing the second hand moving. When we see the second hand moving, we experience succession.

To see why these are not necessarily identical experiences, imagine a subject whose experience of the second hand is like our experience of the bullet. The second hand moves too fast for them to have an experience of succession. All sensory systems have what Ernst Pöppel calls a 'window of simultaneity', an interval during which temporally distinct events that happen successively are fused so that what we perceive is a single event.[18] In normal humans, distinct visual events that are presented with a gap (a 'stimulus onset asynchrony') of approximately 20ms or less are fused in this way. Let us suppose that the individual we are imagining has a window of

[16] W. James, *The Principles of Psychology*, Vol. 2 (London: MacMillan, 1890/1950), 629. D. Zahavi ('Perception of Duration Presupposes Duration of Perception – or Does it? Husserl and Dainton on Time', *International Journal of Philosophical Studies* **15:3** (2007), 464) attributes a similar distinction to Alexius Meinong. It can of course be traced back further to Kant's discussion of the experience of change and persistence in the Second Analogy.

[17] J. Locke, *An Essay Concerning Human Understanding*, edited by P. Nidditch (Oxford: Oxford University Press, 1975), Book 2, ch.14, §10, quoted by S. Kelly, 'The Puzzle of Temporal Experience' in A. Brook and K. Akins (eds), *Cognition and the Brain: the Philosophy and Neuroscience Movement* (Cambridge: Cambridge University Press, 2005), 216.

[18] E. Pöppel, *Mindworks: Time and Conscious Experience* (Boston: Harcourt Brace Jovonavich, 1988), ch.2.

simultaneity that is much larger so that any two events that are separated by intervals of a second or less are fused. Call this subject SLOW. We can suppose that SLOW has a perceptual system that produces representations of the position of the second hand at each second, but SLOW isn't capable of seeing the movement of the second hand from one position to the next. His perceptual system fuses the event of the hand moving from one position to the next. Thus SLOW undergoes a succession of perceptual states but he lacks an experience of succession when the second hand moves from one position to the next.[19] SLOW doesn't experience succession because he doesn't experience movement. To experience movement is just to experience an object at a succession of different positions at different times.[20] To experience movement is to experience succession.

Now imagine a second subject with the opposite problem. Call this subject FAST. For FAST the second hand moves too slowly for him to be aware of it changing position from moment to moment. We can imagine that FAST has a window of simultaneity that is tiny compared to ours. He can distinguish between as many events in a second as we can distinguish in an hour. So for FAST, watching the second hand move around the face of the watch is indistinguishable from what it would be like for us to watch the hour hand moving round the face of the watch. FAST discerns no change of position. Again, FAST can be said to undergo a succession of experiences, but he doesn't experience succession.

At first glance the cases of FAST and SLOW would seem to be easily accommodated by a memory account. The memory account will say a person can see the motion of the second hand if at each moment when she sees the position of the second hand she can also *recall* where the second hand has just been. FAST and SLOW fail

[19] Someone might question whether SLOW really can be said to have a succession of distinct experiences given that events lasting less than a second are fused by his perceptual system. If we individuate experiences by their objects, surely we have to say that SLOW can undergo at most a single experience per second. I agree, but the way I set up the example SLOW has a perceptual system that can detect *non-consciously* the different positions of the second hand at intervals of less than a second. However the processing that results in his experience is less discriminating. SLOW undergoes a succession of distinct perceptions but because of his enlarged window of simultaneity he doesn't get to experience succession.

[20] I am assuming here a view of motion, which Bertrand Russell dubbed the 'at-at' theory in his *Principles of Mathematics* (New York: W. W. Norton & Co., 1903). According to the at-at theory an object moves when it exists *at* a unique position *at* different times.

to see movement because at the time when the second hand changes position they can't recall its previous position. SLOW fails to recall the second hand's previous position because the events of the second hand first being at one position and then being at the next position aren't discriminated by his perceptual system. At the time when the second hand is at one position, he cannot recall where it has just been. FAST cannot recall where the second hand has just been because the temporal resolution of his experience is so fine-grained that by the time the second hand changes position too much has happened for his to recall where the second hand has just been.

While this certainly seems a promising line to take, still a puzzle remains. How can an experience combined with a co-occurring memory at a single instant in time make us aware of events that seem to be spread out across time? This trick must somehow be accomplished by the way in which sensory memories combine or are unified with co-occurring sensory experiences. When a sensory experience is added to a sensory memory the result must be that a subject can experience what seems to be an event, or a series of events, spread out through time. How could an experience and a memory taking place at a single instant in time make it seem to a subject as if she experiences events extended in time? How can a single experience present successive states of affairs as successive? If successive states of affairs were present in a single complex experience surely all we would see would be a blur, much as in a long exposure photograph taken with a slow shutter-speed. In a long exposure photograph a scene is captured in a single picture as it unfolds over time. All that we can see clearly are the static parts of the scene, and the changes are depicted as blurs. This is not what our ordinary visual experience of motion is like. We can normally see moving objects in sharp focus.

Let us suppose the contribution of memory to experience does indeed make it the case that an experience is no longer confined to a single durationless instant.[21] Once this has been conceded, it is hard to see just what it is the memory account is claiming the folk are mistaken about. Haven't we conceded that we can indeed experience events spread out through time just as the folk claim? If we allow that experiences present us with events that at least appear to have temporal properties, we must give up on the claim that experience is durationless. We must say instead that we can be conscious of

[21] I. Phillips, 'Perceiving Temporal Properties', *op. cit.*, argues for a memory account that denies experience is confined to a durationless point in time.

whole intervals of time. This is to agree with William James that our experiences take place in what James (following E. R. Clay) referred to as the *specious present*. Just as we can experience objects that are spread out across three dimensions, so we can experience events that are spread out through time. We can experience change, persistence, succession, and so on because we can experience whole intervals of time in which events can instantiate these properties. This is to reject a memory account to the extent that such an account is motivated by the claim that all we can experience is what is present to us at a single instant in time. Rejecting such an account leaves us with as many questions as it does answers, not least of all how our experiences could be spread out across time.

4. The Puzzle of Temporal Experience

Prima facie, any account of the specious present would seem to be vulnerable to the very same problem we have just raised for memory accounts. For it is indeed compelling to think that at each point in time what we experience is fully determined by the state of mind we are in that time. If we are to experience change, persistence and succession, we must do so at the time our experience occurs. Izchak Miller dubbed this claim the *Principle of Simultaneous Awareness* (PSA).[22] We can accept that the *contents* of our experiences may not be confined to a single instant in the way Reid argued they were. What we experience from moment to moment must however surely be based on the state of mind we are in at a given point in time. How could a state occurring at a single point in time somehow spread out through time to make us aware of a whole temporal interval?

Perhaps the solution to this difficulty lies in an even more thorough-going rejection of the claim that states of mind occur at particular points in time. Our perplexity about temporal experience is generated by the natural seeming idea that it is an experience occurring at a single instant in time that must somehow embrace change, persistence or succession. If we were to allow that experiences (and not only their contents) extend through time perhaps this would relieve our sense of puzzlement. We could say that when we experience a temporally extended event, there is a succession of perceptual experiences that runs in parallel with the successive phases of this

[22] I. Miller, *Husserl, Perception and Temporal Awareness* (Cambridge, MA: MIT Press, 1984).

event. If I hear a sequence of notes C-D-E, for instance, I undergo acts of awareness that coincide in time with the representation of notes C, D and E. Maybe the solution to our puzzlement lies in signing up to what Miller has called the *Principle of Presentational Concurrence* (PPC). PPC claims that the 'duration of the content being presented is *concurrent* with the duration of the *act* of presenting it'.[23]

Timothy Sprigge seems to have endorsed PPC.[24] He conceived of a person's consciousness as made up of short pulses each of which is a single entity 'whose earlier and later phases are not genuine particular realities'.[25] Within a particular pulse of consciousness we can experience change or movement: we can hear a C followed by an E. While Sprigge allows that pulses have temporal spread, he denies that the different aspects of a pulse are units of experience in their own right. Each pulse is both temporally extended and simple in the sense of not allowing for the possibility of further decomposition. These two properties of pulses seem to imply that the different phases of experience within a given pulse must run concurrently with the contents they present. The phases of a pulse aren't experiences in their own right that could possess temporally extended content. If each pulse can embrace change in the way Sprigge argues they must do so by running concurrently with changing contents.

Sprigge denied that we ever experience the transitions from one pulse to the next:

> …only the transitions within, and not between, distinct pulses of experience can themselves be experienced in immediate fashion. So in a broad sense, the sense of flow is itself, I think, a case of

[23] *Ibid.*, 107. Phillips mentions in passing ('The Temporal Structure of Experience', *op. cit.*, ms, 6) a distinction between a weak and a strong reading of PPC. By the strong reading he seems to have in mind a view that claims the interval of time an experience takes up is the very same interval of time as the event we are experiencing. A weak reading of PPC denies this. Phillips frames PPC in terms of experiences and their objects whereas Miller speaks of acts and contents. Is this just a terminological difference? I'm not sure. It certainly makes sense to think of an experience occupying an interval of time different from that of its object. This kind of mismatch is commonplace due to processing delays. However once we've abandoned the principle of simultaneous awareness, it is not so clear to me how experiences could occupy intervals of time different from their contents.

[24] T. L. S. Sprigge, *The Vindication of Absolute Idealism* (Edinburgh: Edinburgh University Press, 1983), 12–14.

[25] *Ibid.*, 12.

> memory… As I speak, short stretches of what I say will be experienced by me in one pulse of experience, but what precedes this stretch guides me as something vaguely remembered as whence this present me emerged.[26]

There is however something phenomenologically off-key about Sprigge's denial that we can experience transitions between different moments of experience. Sprigge allows that we experience transitions between events when these transitions occur within a single pulse of experience, but he denies we experience such transitions across pulses of experience. Dainton has pointed out that if this were right we should be able to find evidence for these two kinds of transitions when we reflect on our experience.[27] We should be able to differentiate between transitions between events that are sensibly present to us, and transitions that can only be recollected, but this is something we don't seem to be able to find. Consider the following example:

> Move your hand slowly but smoothly across your field of vision. At each moment you see your hand at a different position; you also see your hand continuously moving. Not only is the movement continuous, but your experience of the movement is continuous: you are directly aware of every perceivable change in your hand's location in the same way.[28]

There are no obvious boundaries separating one phase of consciousness from the next in the way we would expect there to be if consciousness were composed of momentary pulses. Based on these sorts of phenomenological considerations, a proponent of PPC ought to hold that experiences are never momentary. Instead different phases of consciousness can *overlap* in a way that forms the basis for the continuity of consciousness. Different phases of consciousness overlap by sharing common parts.[29] I'll follow Dainton in calling this the *overlap account*.

According to the overlap account, we can be directly and sensibly aware of events that take place within a short duration of time lasting perhaps a second or less. By 'directly aware' I mean awareness

[26] *Ibid*., 14.

[27] For a more detailed assessment of Sprigge's view see B. Dainton, *Stream of Consciousness: Unity and Continuity in Conscious Experience, op. cit*., §5.5.

[28] *Ibid*., 129.

[29] *Ibid*., ch.7; J. Foster, 'In *self*-defence' in G. F. MacDonald (ed.), *Perception and Identity* (London: MacMillan, 1979); J. Foster, *The Case for Idealism* (London: Routledge & Kegan Paul, 1982).

without the mediation of any kind of memory or recollection. This interval of time is what we referred to earlier as the *specious present*. Just as with Sprigge's pulse theory, our awareness *extends* a short distance through time. Moreover all events occurring within a specious present are experienced with the same maximal force and vivacity. However unlike the pulse theory, memory isn't introduced to provide links between distinct specious presents. Instead it is claimed that distinct specious presents can come together by sharing common parts. If two events are separated by an interval longer than the duration of a single specious present, this will preclude the sharing of common parts with the consequence that the two events cannot be experienced together. At best we will be able to remember one of the events while experiencing the other. Here is Dainton explaining the basic idea:

> Suppose *do-re* is apprehended by an extended awareness A_1 and *re-mi* is apprehended by another extended awareness A_2. Isn't *re* being experienced twice over? Not if the portion of A_1 which apprehends *re* is numerically identical with the portion of A_2 which apprehends *re*: under these circumstances *re* is experienced just once.[30]

How do events occurring within a single specious present come to be experienced together? When we hear notes C, D and E succeeding one another we hear them together as parts of a single experience. This is part of what it is for us to experience each note as succeeding its predecessor as opposed to merely undergoing a succession of experiences. Bayne and Chalmers introduce the term 'phenomenal unity' to refer to what it is like to occupy two or more conscious states together.[31] Consider the experience of seeing a person's lips move at the same time as hearing what they say. These two experiences are phenomenally unified insofar as we undergo a single state that *subsumes* our individual experiences of seeing the person's lips moving and hearing the words they are saying. When we experience notes C, D and E succeeding one another, I suggest we are also in a single complex state that subsumes our individual experiences of first hearing note C, then D and finally E. There is something it is like to undergo this single complex state that is different from what

[30] B. Dainton, 'The Experience of Time and Change', *Philosophical Compass* **3:4** (2008), 634.

[31] T. Bayne and D. Chalmers, 'What is the Unity of Consciousness?' in A. Cleeremans (ed.), *The Unity of Consciousness: Binding, Integration, Dissociation* (Oxford: Oxford University Press, 2003).

it would be like to undergo each experience individually and separately. This phenomenal difference lies in the relation of phenomenal unity that holds when we experience succession. Phenomenal unity is a relational property of the individual experiences that combine to form the single complex state. Recall again our individuals FAST and SLOW seeing the second hand move around the watch face. Their successive perceptions are not phenomenally unified. The distinct perceptions they undergo of the second hand are not subsumed by a single state, and this is why they do not experience succession.

Dainton has introduced a special kind of experiential relation he calls *co-consciousness* to account for phenomenal unity. Co-consciousness is a relation that binds together a brief phase of consciousness with its adjacent phases so that the different phases are experienced together. The relation co-consciousness holds both at a time and over time. Moreover it is a relation that obtains equally within a single experience and across distinct experiences. All the parts that make up a spatial field of view are, for instance, mutually co-conscious. According to Dainton, the same is true of the parts that make up a temporal field of view where the parts of a temporal field of view will be the phases of experience that make up a given specious present.

Dainton tells us: 'Co-consciousness is a basic experiential relationship, one about which there is nothing more to be said, at least while we confine ourselves to describing how things seem.'[32] His arguments for this conclusion are complex, imaginative and elaborate in ways that I cannot rehearse here. The question I want to press is whether a phenomenological account of temporal experience can satisfactorily settle for the conclusion that co-consciousness (or phenomenal unity) is a basic relation that can't be given any further analysis?

I've argued that the difference between experiencing succession and undergoing a succession of experiences lies in the relation of phenomenal unity. One of the challenges an account of temporal experience faces is to explain phenomenal unity, or so I've been arguing in this section. It is therefore rather unsatisfying to be told that the relation of phenomenal unity is a relation 'about which there is nothing more to be said'. In what remains of this paper I consider whether an intentionalist account of temporal experience might be able to do better.

[32] B. Dainton, *Streams of Consciousness, op. cit.*, 84.

I begin with a version of the intentionalist account recently defended by Michael Tye.[33] Tye argues that the tangles philosophers have got themselves into in thinking about phenomenal unity may be born of flawed assumptions. Once we set aside these mistaken assumptions, Tye suggests a path will open up to a rather simple and elegant account of phenomenal unity.

5. Phenomenal Unity: a Spurious Problem?

I have defined phenomenal unity as a relational property of experiences or phases of experience in virtue of which their contents are *experienced* together. This definition, at least upon first hearing, seems to imply the existence of two levels of experience. On the one hand there is a first-order level of experience at which individual experiences or phases of experience occur that can exist independently and separately from other experiences or phases of experience. A second level of experience is responsible for combining or unifying these different experiences so that there is something it is like to undergo them together.

Tye has argued that an account of phenomenal unity along these lines faces a regress problem.[34] Consider again our auditory experience of notes C, D, and E. I have argued there is something it is like to experience the notes together that is different from what it is like to experience them separately. Suppose there is a second-order experience we undergo when we are aware of the contents of these individual experiences together. Now we have our original three experiences of notes C, D and E plus a second-order experience, which is supposed to make it the case we experience the notes together. Surely this just leaves us back where we started. Do we not need a further experience that unifies our original three experiences and the second-order experience of unity? Tye summarises the difficulty as follows:

> It seems that there must be a further unifying relation that binds the experiences. This relation, however, must itself be

[33] M. Tye, *Consciousness and Persons: Unity and Identity* (Cambridge, MA: MIT Press, 2003).

[34] The objection is similar to S. Hurley's 'just-more-content' objection to what she labels 'subjective' accounts of the unity of consciousness. See S. Hurley, *Consciousness in Action* (Harvard: Harvard University Press, 1998), ch.2.

> experienced. For the unity is phenomenal. And now a regress has begun to which there is no end.[35]

Even supposing there was some way around the regress problem, Tye claims that any account of phenomenal unity in terms of an experience of unity will face a further serious difficulty. When we introspect our auditory experience of C-D-E, we don't seem to find anything like an experience of unity. Indeed Tye insists that we don't find anything at all over and above the notes C-D-E. When we introspect, our experience is like an invisible window onto the world. We look right through the window to the scene beyond. Of course, in reality windows aren't invisible and we can turn our attention to them, so the analogy isn't entirely apt. When we introspect our experience, the situation is supposed to be different. Experiences are supposed to figure in introspection like panes of glass that we just couldn't notice no matter how hard we looked.[36]

If we are not aware of our experiences when we introspect, it follows *a fortiori* that we couldn't be aware of our experiences as unified through introspection. Introspection grounds claims about what it is like to undergo a particular experience. The account of phenomenal unity we are currently considering introduces an experience of unity to account for what it is like to experience contents together. However we've just seen that introspection reveals no overarching experience of unity. What grounds do we have for positing such an experience of unity in our account of phenomenal unity if we find no such experience when we introspect?

In order to escape these two problems Tye is led to endorse what he calls the *one experience view* (OEV), an ancestor of which was advanced by Carnap.[37] Consider the experience of a wine taster. He sees the colour of wine; takes hold of the glass and can feel it in his hand; swishes the wine around the glass and hears the pleasant sound it makes against the side of the glass, smells the wine, and finally tastes it moving it around in his mouth. Moreover, let us suppose the wine taster is in a hurry so that he manages to have all these experiences together in a single specious present. OEV claims that the wine taster doesn't have five separate modality-specific experiences one after the other. Rather he has a single complex multimodal experience that can be described in more or less complete ways by describing each of the modality-specific experiences and the

[35] M. Tye, *Consciousness and Persons, op. cit.*, 22.
[36] *Ibid.*, 23–24.
[37] R. Carnap, *The Logical Structure of the World* (Berkeley: University of California Press, 1967).

properties they represent.[38] When we experience different contents together this is not because there are different experiences with their own individual contents that must somehow be combined. Instead there is a single experience with a multi-faceted complex content that represents everything we are conscious of, from the moment we wake to the moment we fall into a dreamless sleep. We can distinguish within this single experience what it is like to undergo this experience at particular moments in time. In doing so we are distinguishing moments or 'stages' belonging to a single extended experience, not individual token experiences.

How does the OEV account for phenomenal unity? First off, it claims that phenomenal unity is a 'relation between qualities represented in experience' and not between experiences.[39] The OEV denies that there exists a plurality of experiences from moment to moment that stand in need of unification. Unity instead is a relational property that attaches to the qualities experienced. It is the things and qualities we experience at different times that are experienced as continuing on from, or succeeding one another:

> If I experience a loud, high-pitched sound, it is the auditory qualities of the sound that are experienced as continuing from moment to moment; if I feel a pain in a thumb, it is the changing qualities of the disturbance I experience in my thumb that are experienced by me, as the pain starts to throb and intensify. Continuity, change and succession are experienced as features of items experienced, not as features of experience.[40]

Tye distinguishes between two types of phenomenal unity, which he calls *direct* and *indirect* respectively. Direct phenomenal unity attaches to consciousness when the qualities that are represented in one specious present are experienced as 'succeeding or continuing on from the qualities experienced in the immediately prior specious present'.[41] Indirect phenomenal unity obtains when the qualities represented in 'nonadjacent specious presents are linked by chains of direct phenomenal unity'.[42] Both kinds of phenomenal unity describe relations that hold between specious presents. What about phenomenal unity within a particular specious present?

[38] M. Tye, *Consciousness and Persons*, *op. cit.*, 28
[39] *Ibid.*, 36.
[40] *Ibid.*, 97.
[41] *Ibid.*, 100.
[42] *Ibid.*

The OEV allows that a single experience has stages, and that each stage has a phenomenal character. What it denies is that experience stages are single experiences in their own right as opposed to parts of a single experience. However we can ask, for each stage of a single experience, in virtue of what is this stage phenomenally unified? The OEV cannot deny that from moment to moment the qualities that are represented at each time are experienced together. What, according to the OEV, is the difference between undergoing a succession of experiences and experiencing succession? Tye's notions of direct and indirect phenomenal unity don't seem to help us much with this question. Nor does it help to be told: 'phenomenal succession is a relation between qualities that enter into the phenomenal content of the single experience of the subject for the given period of consciousness'.[43] We knew there was *some kind* of relation between qualities that obtains when we experience succession, and that doesn't when we simply undergo a succession of experiences. What we were looking for from an account of phenomenal unity was an analysis of this relation. Tye's OEV doesn't seem to be much of an improvement on the overlap account in this regard.

Tye might reasonably respond to this objection that it is illicitly trading on the very view of experience that the OEV means to reject. It is guilty of thinking of stages of experience on the model of token experiences that can be unified both at a time and over time. The OEV is however committed to denying that perceptions and experiences succeed one another. Tye writes: 'A succession of feelings is indeed not a feeling of succession. But a feeling of succession is not a feeling of the succession of *feelings* either'.[44] The single experience that persists for a given period exhibits phenomenal unity, and does so by virtue of the relations of direct and indirect phenomenal unity. Certainly we can give all sorts of incomplete descriptions of this experience in terms of the relations that hold between the qualities experienced at different times. These descriptions however will exhaust what there is to say about phenomenal unity within a specious present.

Such a response depends upon Tye's commitment to a position I'll label *strong intentionalism*. A strong intentionalist identifies the phenomenal character of an experience with *pure representational properties*. A pure representational property is the property of having a certain intentional content or representing things as being

[43] *Ibid*., 101.
[44] *Ibid*., 102.

a certain way.[45] When we experience qualities together, these qualities are not qualities of experience. Rather they are qualities that belong to the things we experience, and that enter into an experience's intentional content. Do pure representational properties exhaust the phenomenal character of an experience?

Different states of mind can be directed upon one and the same object. We can remember or imagine tasting a particular fine bottle of cabernet, but we can also see, smell and taste this same cabernet. In all of these cases we are undergoing different experiences that have an intentional content directed upon the very same object. The difference lies in the experience's mode or manner of representation. Each sense-modality has its own distinctive manner of representation. In addition to the difference between the senses, there are also significant differences between the manner of representation characteristic of imagining or remembering, and that of the senses. Finally there are doxastic (belief-like) and non-doxastic (perception-like) manners of representation.

There is a lot more to be said about each of these manners of representation, and how to go about individuating them. What I have said, however, suffices to illustrate an important point that isn't, I think, easily accounted for by strong intentionalism. Whenever there is an object that the mind is directed upon, this object will always be represented in a particular manner or under a particular mode. For this reason we cannot identify the phenomenal character of an experience with its pure representational properties. In order to fully characterise an experience's phenomenal character we need to mention both its intentional content and its manner of representation. Chalmers calls this position *impure representationalism*.[46] I'll refer to it as *weak* intentionalism, so as to secure the contrast with what I've called strong intentionalism of the kind endorsed by Tye.

I admit I've provided nothing in the way of an argument for the weak intentionalist view I favour over Tye's strong intentionalism. It is possible that the differences I've characterised as differences in manner of representation could be explained as differences in pure representational properties. I won't attempt to argue against this possibility here, since this paper is not intended to be a defence of weak intentionalism in general. In what remains of this section I'm going to simply assume the truth of weak intentionalism and ask where this leaves the OEV?

[45] D. Chalmers, 'The Representational Character of Experience', *op. cit.*, 155.
[46] *Ibid.*

What makes weak intentionalism attractive is that it opens up the possibility of providing an account of phenomenal unity within a specious present in a way that the OEV fails to. The strong intentionalist argues that phenomenal unity can be fully explained in terms of the relations between the properties of external objects that enter into the contents of experience. The weak intentionalist sees another possibility: phenomenal unity may be something that can be explained in terms of an experience's manner of representation. It is this possibility that I take up in the final section.

6. The Weak Intentionalist Account

Consider the following vivid example I've heard Dan Lloyd use in a talk. Imagine hearing a familiar song such as the Beatles 'Hey Jude'. As you hear the opening lines you are already anticipating what is about to come next. You don't just hear 'Hey Jude…' by itself. At the very same time you hear this line, you are also anticipating hearing '…don't make it bad'. What you anticipate determines in part the character of your experience now. 'Hey Jude…' wouldn't sound like it does now if you didn't, at the time of hearing it, also anticipate hearing at the next moment '…don't make it bad'. Your anticipating what you are about to hear is, I propose, a part of your auditory experience's manner of representation. It is a part of your state of mind that is directed upon a particular intentional object. Correlated with your anticipation is a forward-looking dimension to the content of your experience: what you anticipate is what you are about to hear next.

Experiences also have a backward-looking dimension to their content, and they do so, I claim, because of their manner of representation. Consider G. F. Stout's example of hearing the postman's 'rat-tat' at the door.[47] By the time you hear the 'tat' the postman's 'rat' is in the past. Yet the current 'tat' you are hearing wouldn't quite appear as

[47] G. F. Stout, 'Perception of Change and Duration' in his *Studies in Philosophy and Psychology* (London: MacMillan, 1930). Phillips offers a useful discussion of this example (see his 'Perceiving Temporal Properties', *op. cit.*). There are however important differences between the weak intentionalist account I favour and the version of the memory account Phillips defends. We'll see later that I reject PPC, while Phillips endorses it. Ultimately I think the differences between our views of temporal experience will turn on Phillips' commitment to naïve realism, and in future work I plan to tackle this issue further.

it does had it not just been preceded by a 'rat'. You don't just hear the 'tat' sound in isolation all by itself; you hear it as following on from the 'rat'. You can hear the 'tat' as following on from the 'rat' because your auditory experience includes as part of its manner of representation a constituent that keeps hold of experiences from the recent past for a short duration of time. Again this constituent is a part of your state of mind that is directed upon a particular intentional object. In this case the intentional object is the sound you've just heard.

The view I've just outlined is based on ideas that can be found in Husserl's lectures on internal time consciousness dating from 1893–1917. Husserl called the forward-reaching (anticipatory) aspect of an experience's manner of representation, *protention*, and the backward-reaching aspect *retention*. He argued that every conscious episode includes as a part of its manner of representation a retentional component that keeps hold of experiences from the recent past, and a protentional component that anticipates what might be experienced next. The contents of consciousness are being continually renewed at each instant as the perceptual systems receive new sensory inputs from the world. However, experience can encompass more than what we sense from moment to moment.[48] Our experience is never exhausted by sensory impressions (what Husserl called *primal impressions*) that pertain to a single instant in time as Reid argued. Sensory impressions never occur in isolation, but always within a broader temporal context of what has been retained from the recent past, and what is anticipated about the near future.

The weak intentionalist account differs from memory accounts insofar as it denies that experiences are punctuate. This, it seems to me, undermines the motivation for a memory account, since the latter tend to be explicitly premised on a view of experience as punctuate.[49] There are also important phenomenological differences between retention and memory. William James captures one such difference when he tells us that an 'object which is recollected, in the proper sense of that term, is one which is absent from

[48] Husserl took this account of time consciousness to apply to consciousness in general including conscious thoughts, and other kinds of mental acts. My concern in this essay is with temporal experience, hence my concentrating on the experience case.

[49] See I. Phillips, 'Perceiving Temporal Properties', *op. cit.*, for an important exception.

consciousness altogether, and now revives anew'.[50] Retention is by contrast introduced in order to explain how our experiences of now can extend beyond a point in time to embrace whole intervals of time of a short duration. When a note E is presented in experience as following on from C, note C forms a part of what we are hearing in the present, but it does so as something that is just past. C isn't absent from consciousness in a way that requires it to be dredged up from memory. Rather C is still present in consciousness but it is present as an event that has just passed.

Husserl is commonly represented as subscribing to the principle of simultaneous awareness (PSA): the claim that experiences are confined to single points in time, but they possess contents that are temporally extended. I agree with Shaun Gallagher that Husserl probably would have endorsed *a* version of PSA: he would have allowed that in order for a subject to experience succession, continuity and change she must be continuously aware of at least some part of a temporally extended event at any given instant.[51] A subject can be continuously aware of a temporally extended event because each experience includes, as a part of its manner of representation, retention and protention. We can find within any episode of experience a tripartite structure composed of a sensory impression, retention and protention. This tripartite structure makes it the case that every event will appear in the context of a recently experienced past and an anticipated future.

[50] W. James, *The Principles of Psychology*, *op. cit.*, 646. In a more recent discussion of memory, M. G. F. Martin has argued that what different kinds of memory (*e.g.* semantic and episodic memory) have in common is that they preserve cognitive contact with the past ('Out of the Past: Episodic Recall as Retained Acquaintance' in C. Hoerl and T. McCormack (eds), *Time and Memory: Issues in Philosophy and Psychology* (Oxford: Oxford University Press, 2001)) Martin's account of memory suggests a way in which retention might be understood as a variety of memory. Retention preserves experiential contact with the recent past, and if we think of experiential contact as a kind of cognitive contact, retention ends up counting as a kind of memory. This doesn't however suffice to make the Husserlian view into a memory account of temporal experience. For the memory account recall is motivated by the claim that the contents of experience are punctuate, whereas we have just seen that the Husserlian view claims that a sensory impression never occurs in isolation from retention and protention.

[51] S. Gallagher, 'Sync-Ing in the Stream of Experience', *Psyche* (2003), 4. (Available at http://journalpsyche.org/ojs-2.2/index.php/psyche/issue/view/127).

Weak intentionalism must however reject the principle of presentational concurrence (PPC), the claim that experiences are temporally extended in a way that allows for experiences to coincide in time with their contents. Retention and protention are not a part of what we are aware of in experience. We are aware of the contents of retention and protention. Retention presents a note we have just heard as recently past. It follows that there is from moment to moment an aspect of our awareness, namely retention, that doesn't coincide in time with the content of our awareness. The retentional component forms a part of ongoing experience at each moment, but what is retained is an event or a part of an event represented as just past.

Weak intentionalism would reject PPC for an additional reason. According to PPC events occupying an interval of time of up to a second can be sensibly present to a subject of experience. When we hear the postman's 'rat-tat' at the door, both sounds are sensibly present to the subject in a single auditory experience. At the time the 'tat' is heard the 'rat' is still sensibly present in experience. The weak intentionalist by contrast holds that the recent past is present in experience *intentionally*. Whenever a subject undergoes a sensation there is something present to the mind to which the subject stands in a relation. This is of course a standard claim that is made in discussions of sense data; Howard Robinson has called it the *phenomenal principle*.[52] Intentionalists reject the phenomenal principle: a state of mind S can be intentionally directed upon an object X even when X does not exist. Relations however entail the existence of their relata. It follows that intentionality cannot be conceived of as a relation that holds between a person and the object of their state of mind. If intentionality is to be understood as a relation at all it must be a relation that holds between a person and an intentional object, where an intentional object is conceived of as the content of an intentional state.[53]

Now consider how these familiar points apply to the case in hand. When retention keeps hold of some event from the recent past – the note that has just sounded in a melody – the subject doesn't continue to stand in a relation to this event. The event now forms a part of the past. It doesn't linger in consciousness in the form of a sensation. If it did, we would never hear melodies and only ever hear chords as we do

[52] H. Robinson, *Perception* (London: Routledge, 1994).

[53] This leaves intentionalists needing to provide some account of the relation between an intentional content and the object this content is directed upon. For some promising suggestions in response to this problem see Tim Crane, 'Intentionalism', *op. cit.*

when distinct notes are played simultaneously. We don't continue to hear the previous note. Rather I will argue it is the *sense* of the event that is just past (*i.e.* the note that has just sounded) that is retained.

We can think of the intentional content that is correlated with retention and protention in a Fregean way, as a mode of presentation.[54] The thoughts 'Hesperus is shining' and 'Phosphorus is shining' are thoughts with different senses that refer to one and the same state of affairs. They differ in sense by differing in the mode of presentation under which they present Venus. In a similar way the same note in a melody can be present in experience in different ways at different times. When it is first heard, it is present under the mode of presentation now. When it is retained, it is present in experience as just past. As the note recedes further into the past it continues to be present in experience, but it does under a changing mode of presentation until it is no longer present at all. The contents of retentions and primal impressions are functioning here as different modes of presentation that refer to one and the same note.

Events or parts of events that are located in the past, present and future, are present in experience in different ways. This marks another important difference between the intentionalist account I favour, and views of temporal experience that endorse PPC. Proponents of PPC hold that events that are past can be present in experience with just the same degree of vividness as events that are happening now. Intentionalism by contrast holds that the recent past is present in experience in a different way from events that are happening now. Events that are recently in the past are intentionally present in experience as past. They are present under a different mode of presentation from events that are present in experience as happening now.[55] I conclude then that weak intentionalism is multiply inconsistent with PPC.

[54] Here I am following a suggestive comment made by Shaun Gallagher. He writes: 'retention takes the just-past as a semantic referent' (*The Inordinance of Time* (Evanston: Northwestern University Press, 1998), 48). He seems clearly to be thinking of retention as functioning in the same way as modes of presentation in thoughts determining reference and making it the case that co-referring thoughts have different cognitive values. I say a little more about this above.

[55] B. Dainton in his 'Sensing Change', *op. cit.*, wonders whether our perceptual discrimination is sufficiently fine-grained to allow us to 'discern different degrees of presence (or pastness) over periods of a second or so' (374). This is a reasonable question to ask if we take different sensory modes of presentation to involve sense-data that are present with different degrees of force and vivacity. However I've argued that we

I'll end by briefly explaining how weak intentionalism might account for phenomenal unity. If I'm right, this gives weak intentionalism an advantage over rival theories of temporal experience. I've argued above that a number of theories of temporal experience that are currently on the market fail to tell us anything informative about phenomenal unity.[56] Since phenomenal unity is a core feature of temporal experience, these rival theories fail as accounts of temporal experience.

Retention and protention serve a dual purpose in a weak intentionalist account of temporal experience. They explain how we can be aware of temporally extended events, but also how our experiences come to be unified over time. What is retained isn't only the event from the recent past, but my experience of this event. In this way each phase of consciousness includes the retention of a previous phase. The previous phase will also have as part of its manner of representation a retentional component. Thus at each moment of experience a continuum of retentions will exist that make me aware both of a temporally extended event and of my ongoing experience. Something analogous is also true of protention. I don't just anticipate what sounds are going to come next in the familiar tune I'm listening too. I anticipate that I'm about to hear those sounds. What I anticipate is something about the future course of my experience, so that at the same time as I am aware of events that extend into the future I am also aware of my experience as it unfolds in the future.

I've characterised phenomenal unity as a property of experience whereby different contents are experienced together. If weak intentionalism is correct, and experiences do include retentions and protentions, we have an explanation ready to hand as to how different contents could be experienced together. Different contents come to be experienced together in a single state of mind because of retentions

---

shouldn't think of temporally extended events as present in experience in the way that sense-data are present in experience. Rather we should think of temporally extended events as intentionally present. Understood in this way it isn't so clear to me that Dainton's objection bites, but obviously there is more to be said here.

[56] Akiko Flischut pointed out to me that this may not be true of Ian Phillips' memory account. It could be argued that the relation of 'constitutive dependence' that Phillips takes to hold between different phases of experience may account for phenomenal unity. This is an intriguing suggestion I intend to return to in future work where I engage more fully with Phillips' account.

that stretch back a little way into the past, and protentions that reach a little way into the future.

Isn't such an account vulnerable to a regress of the kind we saw Tye describe in our earlier discussion of phenomenal unity? Haven't I appealed to an overarching ongoing experience that unifies different contents, making it the case that they are experienced together? If so, don't we then need another level of experience to explain how my ongoing experience comes to be unified? The weak intentionalist account isn't vulnerable to this regress objection since it doesn't appeal to two levels of experience. The unity of experience is explained by the continuum of retentions and protentions that form a part of the manner of representation. Manners of representation belong to first-order experiences. An experience unifies itself by means of its manner of representation. It doesn't require an additional level of experience or representation to bring about this unification.

7. Conclusion

Experiences aren't confined to durationless points in time as Reid argued, but can stretch across time to embrace temporally extended events. I've argued that our experiences can extend through time in this way because of their manner of representation. Each experience has a complex manner of representation that includes a component directed upon the recent past and a component directed upon the impending future. I've arrived at this view in an attempt to do justice to the commonsense view that we can see movement and hear sounds, where experiences of this type take a certain amount of time to unfold. If weak intentionalism succeeds, the folk will stand vindicated. Many questions still remain, not least of all a whole battery of difficult metaphysical questions. There are questions for instance about the relationship between our experience of temporal properties and the metaphysics of time. Does weak intentionalism have any implications for the debate between A-theorists about time and B-theorists?

More pressing still for the arguments of this paper are questions about the relationship between the phenomenological claims of the kind I've been discussing, and the underlying neural and computational mechanisms that support experience. If experiences do indeed have complex manners of representation of the kind I've posited, how is this to be accounted for at the level of neural and computational mechanisms? I've purposefully remained at the level of

phenomenology in this essay in an attempt to show why the folk may not be as confused about the character of their experiences as some philosophers have argued. I've told a story about what experience must be like if we are to experience temporally extended events. However once psychology and neuroscience have told the full story about the mechanisms that underpin conscious experience, couldn't it turn out that the folk are mistaken after all? Couldn't a science of the mind discover that nothing in the brain answers to the descriptions of experience I've given? The claims that are made by phenomenology are certainly answerable to what science tells us about the nature of the mind. Nevertheless, I would claim that a science of the mind, at least to the extent that it is concerned with consciousness, ought to also answer to phenomenology. Scientific discoveries can certainly lead us to refine our phenomenological descriptions, but they cannot lead us to abandon these descriptions.[57]

University of Edinburgh

[57] I am extremely grateful to Barry Dainton, Tim Crane, Eduard Marbach, Andy Clark, Mike Wheeler, Fred Adams, Ken Aizawa, Howard Robinson, Peter Simons, Pierfrancesco Basile, Leemon McHenry, and David Cockburn for discussion of an earlier version of this paper. I've also received immensely helpful feedback from Christoph Hoerl, Ian Phillips, Simon Prosser, Akiko Frischut and Jiri Benovsky at the EIDOS conference on the Experience of Time hosted by the University of Geneva. Thanks also to my colleagues from the Subjective Time group for discussion of ancestors of the ideas in this paper.

Time in Consciousness, Consciousness in Time

DAVID COCKBURN

Abstract
The paper is a criticism of the idea that a notion of 'phenomenal consciousness' has a significant role to play in the attempt to understand how the experience of change is possible. Discussion of such experience must give a significant place to its public and private manifestations. How should we picture the relationship between the experience of change and its manifestations? While we cannot *identify* these, we need not conclude that 'the seeing or hearing itself' is something distinct from – something that has a nature that may be investigated quite independently of – any of its public or private manifestations. With that, we cannot grasp how time can be present in consciousness without reference to the fact that consciousness is located in time.

1. The Source of Temporal Ideas

A philosophical interest in our experience of time may arise from a concern with how we come by temporal ideas: the idea of the *past*, or of one thing happening *before* another. Thus, William James asks: 'What is the original of our experience of pastness, from whence we get the meaning of the term?'[1] He replies: 'the original paragon and prototype of all conceived times is the specious present, the short duration of which we are immediately and incessantly sensible'. Again: '[T]he specious present, varying in length from a few seconds to probably not more than a minute, is the original intuition of time'. Two contemporary analytic philosophers who follow James in aspirations of this form are John McDowell and Hugh Mellor. Thus, McDowell suggests that, since its having rained yesterday 'can be, in memory, immediately available to consciousness', the child's grasp of a conception of the past – of the meaning of the past tense – can be 'drawn from actual confrontation with instances of the sort of circumstance involved'.[2] Similarly, Mellor suggests that the notions of before and after can 'be defined

[1] William James, *The Principles of Psychology*, Vol. 1 (Cambridge, Mass.: Harvard University Press, 1981), 605. Following quotations: 631, 642.

[2] John McDowell, 'On "The Reality of the Past"' in C. Hookway and P. Pettit (eds), *Action and Interpretation* (Cambridge: Cambridge University Press, 1978), 140.

doi:10.1017/S1358246110000068

directly in terms of observable and tenseless temporal relations between events'.[3] One might think that we will only see sense in such suggestions in so far as we are wedded to another strand of traditional empiricist thinking: in so far, that is, as we conceive of having a conception of the past or of temporal priority in terms of a capacity to generate certain forms of mental imagery. And one might think that there is room for serious doubts about *that* aspect of empiricist thought.

I mention this partly in order to mark this discussion off from the issue that is my concern in this paper, and partly to note an analogy between the terms in which these different issues have been addressed. Whether any sense can be made of talk of ideas being 'drawn from' experience, there is, it seems, another philosophical task here: that of understanding how it is possible for our experience to have the features that it does. My concern in this paper is with the suggestion that a notion of 'phenomenal consciousness' has a crucial role to play here: a crucial role to play, in particular, in the attempt to understand how the experience of change is possible. In the position that it gives to mental imagery, or, more broadly, to an idea of 'what is present to consciousness', in its discussion of the experience of change, this suggestion does, I believe, display a significant analogy with empiricist treatments of concept formation.

2. Time in Consciousness

To appreciate that there is philosophical work to be done here we must recognise that experiencing change is something quite different from changing experience. The distinction, and the puzzle that the former gives rise to, is brought to life through reflection on examples of the following kind. We *see* the movement of the second hand of a clock in a sense in which we do not see the movement of the hour hand – even though we can recognise, on the basis of what we do see, *that* it has moved. Again, hearing a melody[4] is not simply a matter of being aware, successively, of a series of notes: it involves

[3] Hugh Mellor, *Real Time* (Cambridge: Cambridge University Press, 1981), 28.

[4] Perhaps we should add: in the sense that is relevant to this discussion. There is a sense of 'hear' such that we can say of someone that she has heard an extended sequence, say Beethoven's Ninth Symphony, simply in virtue of the fact that she was present while it was being played and there was nothing wrong with her hearing.

hearing the succession. And we do not get from one to the other simply by adding memory: when someone hears a melody the situation is not adequately described in terms of her hearing the note that is being played now and *remembering* the sequence that immediately preceded it (along, perhaps, with anticipating the sequence that will follow).

Change and persistence are directly experienced. Change and persistence both take time. For it to be the case that she is hearing the melody there must, it seems, be some sense in which the listener is aware, *in a single act of consciousness*, of notes that are extended over a period of time. But how can that be? Anything that is extended over time has parts that are not strictly speaking happening now. Given, as one might reasonably suppose, that our immediate experience is confined to the present, how is it possible for us to be immediately aware – aware in the same way as we are aware of a pattern of colour extended in space – of phenomena that are not (strictly speaking) present?[5]

That is the puzzle. Different accounts of the temporal character of phenomenal consciousness provide different answers to the puzzle. On one view, while an individual act of awareness has no temporal duration what is experienced *in* such an act includes, not simply things happening at that instant, but also things occurring at a series of immediately preceding instants. In this sense, my experience is not confined to the present instant; I have direct awareness, at an instant, of a process that is stretched over time. In James' well-known articulation of the notion of the specious present: 'the practically cognized present is no knife-edge, but a saddle-back, with a certain breadth of its own'.[6]

In contrast with this we have views, of the kind defended by Timothy Sprigge[7] and more recently in much greater detail by

[5] My formulation draws on the work of a number of philosophers. See, in particular, Barry Dainton, *Stream of Consciousness: Unity and Continuity in Conscious Experience* (London: Routledge, 2000), 114–15; also Dan Zahavi, *Husserl's Phenomenology* (Stanford: Stanford University Press, 2003), 82.

[6] William James, *The Principles of Psychology*, *op. cit.*, note 1, 574.

[7] '[A] single momentary state of consciousness is not a composite of its elements; these are simply conceptually abstracted aspects of its unitary being, not distinct units in their own right. Such single momentary centres of experience, as I shall call them, should not be thought of as instantaneous. So far as they can be clocked in public time they have a certain dateable duration, but they are single units whose earlier and later phases are not genuine particular realities in the way that they are. They do have, within them, their own special kind of temporal spread, aspects which are later

Barry Dainton,[8] according to which individual experiences have a certain timeable duration. An awareness of a temporally extended process is itself temporally extended. An experience of, say, a piece of music that lasts five minutes, is composed of a series of overlapping acts of awareness ('phenomenal presents') each of brief duration (perhaps a second or two). The duration of such individual acts of awareness is to be determined, perhaps, by a combination of introspection and psychological experiment. Dainton suggests that this account both explains how experience of change is possible, and, on reflection, can be seen to be true to the character of our experience.

The dispute between such views is a dispute about the character of what is spoken of as 'phenomenal consciousness'. Particular positions are defended by appeal to what phenomenal consciousness *must* be like if certain familiar features of our experience are to be explained; though, as with Dainton, such appeals may be supplemented with a call to careful self-reflection (in the hope – generally, it has to be said, forlorn in philosophically contentious matters – that the careful self-reflection of others will produce results that match one's own).

It is important here to remember that, while what is at issue are features of *consciousness*, they are features of a kind that are not prominent, or perhaps don't figure at all, in our everyday understanding of ourselves and others. Thus, the philosophical dispute about the individuation of acts of phenomenal consciousness is, at best, tenuously related to our familiar, everyday, ways of individuating experiences. Central to *that* is the way in which we pick out single experiences both at a time and over time in terms of their content. That nasty experience in the bus queue (which lasted about two minutes) was followed by the wonderful experience of sitting upstairs and watching the fields roll by (which lasted a full 15 minutes). Half way through the concert I was overcome with an experience of uncanny familiarity that lasted until it finished.

I will return to the relation between philosophical characterisations of features of 'phenomenal consciousness' and our everyday

---

and earlier. Only thus can one explain the fact that temporal sequences can be given'. (T. L. S. Sprigge, *The Vindication of Absolute Idealism* (Edinburgh: Edinburgh University Press, 1983), 12).

[8] 'We must … abandon the doctrine that an awareness of change (or succession) consists of a temporal spread of content being presented to a single momentary awareness. Instead, we say that an awareness of change or succession is itself temporally extended..... [I]t is natural to suppose that acts have precisely the same duration as their contents'. (B. Dainton, *Time and Space* (Chesham: Acumen, 2001), 103).

individuation of experiences. First it will be helpful to consider further the puzzle that calls for a solution in terms of an account of the character of 'phenomenal consciousness'.

3. Experience ' In' and 'At' an Instant

Philosophical puzzlement is characteristically an expression of an apparent clash between how it seems clear things *must* be and how we find them to be in practice. In our case, the task is to show of something that clearly does happen – hearing a melody, seeing the movement of the arrow - *how it is possible*. Underlying the puzzlement is a contrast between experience of something at an instant, for example the hearing of an individual note, and experience of something over time. The former provides our model of how experience *must* be: a model in the light of which the latter must be, and yet apparently cannot be, understood. Different accounts of the character of phenomenal consciousness are attempts to show that, with appropriate modifications, the model *can* accommodate the experience of change. (This representation of the situation may well be disputed. I hope that what follows will provide some substantiation of it).

If, however, we do think of the puzzle in this way – as a puzzle about how an experience of a melody can be as immediate, as non-inferential, as an experience of an individual note – we should hesitate. For an individual note is extended in time in just the same sense as a melody is. There is no experiencing of a note that is not an experiencing of a temporally extended period of sound any more than there is an experiencing of a colour that is not an experiencing of a spatially extended patch of colour.[9] If we are puzzled about how we can have immediate experience of a melody on the grounds that 'a melody is extended over time and so has parts that are not strictly speaking happening now', we ought to be puzzled in the same way about our experience of an individual note. There is, then, room for a suspicion that our puzzlement about the former reflects a confusion in our understanding of the latter: that is, in the model, which we draw from the latter, of how experience *must* be.

It will be helpful here to place the discussion of our *experience* of change alongside questions that may arise about *change*. Compare the following line of thought: 'We say that the orchestra is now playing Beethoven's Ninth Symphony. Yet the opening movement

[9] See J. Foster, 'In *self*-defence' in G. F. Macdonald (ed.), *Perception and Identity* (London: Macmillan, 1979), 176.

is part of the Ninth Symphony and they are not now playing *that*. How can this be? How can something, a playing of a symphony, that is extended in time be strictly speaking *present* – present in the sense that it must be if they really are playing it now?' We may conclude that, strictly speaking, it is simply not true that they are playing the symphony now. But if we do draw this conclusion on the above grounds we may find ourselves forced to conclude that, strictly speaking, they are not now playing the second movement of Beethoven's Ninth, that they are not now playing B# (since a playing of a note that is over as soon as it starts is not a playing of a note at all), that they are not now playing *anything*.

Perhaps we must conclude that there is something wrong with the idea of 'what happens in an instant' – with the notion of 'present in the strict sense' – that is in play in these discussions. But how are we to say what has gone wrong? We might try saying that the term 'now' always picks out a *stretch* of time, which may be of different lengths depending on our particular interests; and so that the philosophical notion of 'the present in the strict sense' is simply a fantasy. If we do say this, however, we will face the following protest: 'When we speak of what is happening "now" our concern often *is* with the durationless dividing line between past and future. We want to know what he is doing – not throughout the day, nor through this five minute stretch of the morning – but at this instant: four seconds past 10.00. Or: we want to know, not what her average speed was over the hundred yards, but how fast she was moving as she crossed the half way mark. The idea of "what happens in an instant" is not at all a philosophical fantasy'.

The examples bring out that the idea of what is happening *at* an instant is not a philosophical fantasy. Indeed, it is arguable that questions about what is happening 'now' are *always* questions about an instant in this sense. The examples do not, however, show that philosophical talk of what happens '*in*' an instant is completely innocent. *That* notion involves an appeal to the idea that we move closer to a pure description of what is happening now – *in* the present instant – as we move along the series that runs from 'Playing Beethoven's Ninth Symphony' through 'Playing the 4th movement of Beethoven's Ninth' to 'Playing B#' – and perhaps beyond. (We are, perhaps, thinking in some such way when we are tempted to accept Zeno's insistence that at any instant the arrow is at rest: an instant being a time so short – infinitesimally short! – that there is, within it, no room for movement.)

That line of thought – along with the perennial notion of a 'temporal slice' of an object or process, which is a pure expression of

it – embodies serious confusion. If a 'temporal slice' of an object or process is, like a slice of bread, a thing of a kind of which it makes sense to ask how thick it is then, while slices may take less time than symphonies, the idea of playing or hearing a symphony raises no philosophical puzzles of a kind that do not already arise with playing or hearing a slice of a symphony. If, on the other hand, we picture a 'slice' as something to which the notion of 'thickness' has no application (as when we speak of a slice – a cut – through a loaf) then we are speaking just as accurately – just as precisely – of what is happening at a slice when we answer 'They are playing Beethoven's Ninth' as when we answer 'They are playing B#'. They are playing Beethoven's Ninth now if now is an instant in a stretch of time during which they are playing Beethoven's Ninth; and so the answer to the question 'What are they playing *now*, at four seconds past 10.00?' can, given an appropriate context, as readily, and precisely, be 'Beethoven's Ninth Symphony' as 'B#'.

In summary: we may (taking a tip from Augustine) suppose that we are not, in the strict sense, speaking of what is happening *now* when we say 'They are playing Beethoven's Ninth' on the grounds that any playing of a symphony has parts that lie in the past and future: that, strictly speaking, it would only be true that they are now playing Beethoven's Ninth if they were now playing the first movement of the Ninth, along with the Hallelujah Chorus (and every note of those to boot!). But it no more follows from the fact that they are now playing Beethoven's Ninth that they are now playing the opening movement of Beethoven's ninth than it follows from the fact that I am standing on the car that I am standing on the steering wheel.

To see how these considerations might help us in our concerns with the *experience* of change we must turn to the manifestations of such experience.

4. The Manifestations of Experience

We can divide what I will speak of as 'manifestations' of an experience of change into two broad categories: 'public' and 'private'. The former, which encompasses all of the marks by which we may tell that another is hearing a melody or seeing the ball flying towards him, itself forms a fairly mixed bag. It includes expressive behaviour of which the individual may be quite unaware: her tapping out the rhythm, her smile of pleasure at a particular turn in the music or startled reaction when a note is misplayed, and so on. Such expressive

behaviour is on a spectrum with intentional actions grounded in what is observed: she reaches out to catch the approaching ball; she is, on being asked, able to repeat the melody, repeat it not simply in the sense of being able to reproduce the series of notes played but in the sense that we can hear the melody in her reproduction. Still in the public domain there is the range of possible linguistic manifestations of experience: she describes the sequence of events she has just observed; or simply says, in response to our question, 'I can see the movement of the hand'. In addition to such public manifestations, there are what I will speak of as 'private' manifestations: responses – such as the thrill of pleasure that I feel on a startling turn taken at a particular point in the music, or my ability to repeat the melody in my head – that are not, in this instance, expressed in any way detectable by others.

Public manifestations play an important, if sometimes more or less implicit, role in discussions of the experience of change. While introspection, careful reflection on the character of my own experience, is a central feature of such discussions, in so far as the philosopher who is concerned with these issues takes herself, not simply to be investigating a personal peculiarity, but to be writing of, and for, others as well, the public manifestations are inevitably of interest to her. For it is only through those that she has any grounds for supposing that others do have, as she does, experience of change. At a more specific level, such public manifestations play a crucial role in the attempt to determine particular features of the 'acts of consciousness' involved in the experience of change. The temporal stretch of what is experienced at an instant, or, on the Sprigge/Dainton view, the temporal stretch of an experience itself, is to be determined through psychological tests involving, for example, the ability to repeat, or re-identify, sequences of sounds of different lengths.[10]

In practice, appeals in these discussions to public manifestations of the experience of change are primarily to things that people *say*. Thus, Dainton justifies his idea that he is speaking for others too in these terms:

> We unthinkingly assume that our conscious lives are broadly similar, and this unthinking assumption receives confirmation

[10] J. D. Mabbott remarks of such tests: 'The aim was to discover the unit of temporal experience. So far as the experiments discovered anything, they found what was the maximum duration of a set of sounds which could be recognized without error' (J. D. Mabbott, 'Our Direct Experience of Time', *Mind* **LX** (1951), 162–4). His doubts are close in form to those that I will raise.

> over and over again in daily life, in our interactions with others, simply because we rarely (if ever) have occasion to question it. Whenever writers of any kind attempt to describe the general character of human consciousness, whatever the differences in fine details, we can usually recognize them as attempts to describe a mental landscape of a sort we are ourselves acquainted with.[11]

On the face of it, there are, however, no grounds for giving exclusive, or even primary, place to *linguistic* manifestations; still less to descriptions that people offer of their experience. My confidence that my dog perceives change, based on his response to my throwing a ball, is, I take it, as well grounded as my confidence that other people do. And my confidence that people with whom I have contact experience change is grounded most pervasively in their handling of the objects that surround them, in their expressive responses to music, in the smooth flow of their responses to what I do or say in my interactions with them, and so on. We can add that the significance I attach, quite generally, to what a person *tells me* about her experience does not stand alone. Appropriate ties between what she *says about* her experience and other manifestations of it are crucial to my readiness to take her words as I do. To express the point in modest terms: my confidence in accepting another's assurance that she hears a particular melody would be seriously strained if it became apparent that she was, quite generally, incapable of expressing her experience of music in any of the familiar ways. In an extreme case, we would, I take it, have little choice but to say that she simply does not hear melodies; and faced with an apparently sincere insistence on her part that she does, we would, I suppose, have to conclude that she does not have a proper grasp of the sense of those words.

How are we to picture the relationship between the manifestations – private and public, verbal and non-verbal – and 'phenomenal consciousness'? We may picture that relation in causal-explanatory terms. For example, the fact that an individual is, at a given moment, able to introspect, and with that report on, both what he is hearing and what he is seeing, is most plausibly explained on the assumption that these experiences are 'phenomenally unified'.[12] Again, the unusual combination of abilities of those with blindsight is best explained by the assumption that they lack phenomenal

[11] B. Dainton, *Stream of Consciousness*, *op. cit.*, 20.

[12] B. Dainton, 'Coming Together: the Unity of Consciousness' in Susan Schneider and Max Vellmans (eds), *The Blackwell Companion to Consciousness* (Oxford: Blackwell, 2007), 212.

consciousness.[13] And in the case that is my concern in this paper: the best, or only possible, explanation of the fact that a person is able to re-identify a sequence of notes just played, that he moves to the music as he does, or that he is startled when the music takes an unexpected turn, is that his phenomenal consciousness has a certain temporal structure.[14]

If our fundamental grounds for believing that phenomenal consciousness has particular features – or that the phenomenal consciousness of others is of the same general structure as one's own – are to be of this form we are going to need some account of how the postulated acts *would* explain the phenomena: how, for example, the extended single act of awareness of a series of notes, of which Sprigge and Dainton speak, would explain the smile that lights up his face, or the tingle that runs down his spine, in response to a momentary startling turn in the melody.

How are we to conceive of the mechanism by which the previous stages of that act of consciousness bring about the smile or tingle that occurs, let us suppose, at its end? I do not know of attempts to address this question. Yet in the absence of an account of *how* the postulated feature of phenomenal consciousness would explain things like this, there seems to be no reason to suppose that it *would* do so. And in so far as it is its explanatory power that is to provide our reason to postulate a phenomenal consciousness having certain features (or *my* reason to suppose that the phenomenal consciousness of others is of the same general structure as my own), we have *no* reason to do so until that work has been done.

We may fail to feel the force of this point because, as we may think, *that* link – between directly experiencing a sequence of notes and feeling, say, a distinctive twinge of pleasure – is such a completely secure feature of our commonsense understanding of the world as not to call for our attention at this point. But if there is a truth of some kind there, it is hardly one that can carry any weight when we are speaking of the 'acts of phenomenal consciousness' postulated in these accounts. For, as I remarked earlier, *those* – the bursts of two or three seconds that are to be investigated by the psychological tests, and that make possible what, at an everyday level, we speak of as, for example, 'That nasty experience in the bus queue' – are not a

[13] M. Tye, *Consciousness and Persons: Unity and Identity* (MIT Press, 2003), 122.

[14] If we suppose that the individual has direct, and infallible, access to the character of his own phenomenal consciousness it will only be my attributions to *others* that are at issue here.

secure feature of our commonsense understanding. Again, we may picture the explanatory power of the temporal length of an experience through analogy with a physical mechanism: what happens when I actually hear the startling note is explained, in part, by the fact that behind it lies a temporal stretch of experience of which the note is a part – much as what happens at the point of physical impact when the train hits the truck is explained, in part, by the fact that behind the front of the train there is an extended object of which it is part. But I take it that to the extent that anything along these lines is responsible for the appearance of explanatory clarity, what we have is *only* an appearance.

5. The Relation Between Experience and its Manifestations

Any philosophical discussion of our experience of change must, I suggested, give a significant place to *manifestations* of such experience: the smile that lights up his face and the tingle that runs down his spine in response to a momentary startling turn in the melody; his ability to repeat the series of numbers that we have just read to him; and so on. It must do so if only because it is through these alone – in particular, through the public linguistic or behavioural manifestations – that the philosopher has any grounds for supposing that he is speaking for more than simply himself. But it must do so also because such manifestations are central to the *significance* that the experience of change has in our lives. Think, for example, of the place of dancing, or singing along together, in our shared experience of music; or, at a more individual level, of the pleasure someone may take in the silent replaying of a melody in her head. If, in a particular case, an individual displays a rich enough range of such manifestations there may, *in practice*, simply be no room for a serious doubt about whether she really is hearing the melody. With that, it has, perhaps, little plausibility to maintain that: 'What we are really interested in when we ask of someone if she hears the melody is something quite distinct from whether she can sing along with it, dance to its rhythm, take pleasure in it, and so on'. Thus, suppose someone says 'Oh, I can hear the rhythm alright; I just can't get my hand to tap along with it, reproduce it in my head, say whether it is faster or slower than that of the tune played a moment ago …'; or, in response to our puzzlement at his not having ducked when a stone is thrown at him, someone says, 'Oh, I saw the stone flying towards me alright; I just didn't grasp its significance until I felt a sharp bash on the head'. While we may put up with a few responses of

this form, if someone's claims to hear rhythms or see movement are systematically totally disconnected from any manifestations – public or private – in his life we will soon lose interest in his claims. With that, while we should not expect a philosophical account of 'what it is to experience change' to provide us with a story of how *all possible* manifestations are connected with the experience, if we cannot see in it some plausible picture of connections in the most familiar and pervasive cases, we will lack an important foothold for the claim that what has been identified (that is, certain goings on in phenomenal consciousness) is correctly characterized as 'the experience of change'.

Just *how* should we picture the relationship between the experience of change and its manifestations? We may seem to be confronted with a choice. We may hold that the experience of change is something quite distinct from these manifestations. It is an incident in 'phenomenal consciousness' that underlies and explains all of the manifestations of which I have spoken: an incident whose precise character is to be determined by a combination of philosophical reasoning and psychological experiment. Alternatively, we may *identify* the two: to experience change – to hear the melody for example – simply *is* to display such manifestations: to smile, tingle, tap one's foot, and so on.[15] Having noted problems with the first answer, it may seem that I am clearly committed to the second.

Here is one reason we may give for judging that the second answer cannot possibly be acceptable, so that we must return to something of the form of the first: 'We cannot say that the manifestations are what it *is* to experience change, for whatever manifestations he displays it is intelligible to suppose that he is not experiencing change. It is intelligible to suppose that, while he ducked just as the stone was approaching his head, or tapped his foot in synchronization with the rhythm of the music, this was pure coincidence: the ducking or tapping were not expressive of any such experience. The experience of change must, then, be something other than the ducking or the tapping. There must be an underlying act of consciousness that encompasses an awareness of notes extended over time: "must be" – not because the foot tappings could not be explained in any other way – but because it is only in so far as they are explained in *this* way that they are correctly taken to be expressive of an experience

[15] Such an identity claim could come in various forms. In particular, it could come in either a 'type' or a 'token' form. My brief remarks on the identity view will not, however, turn on such possible differences.

of the melody. The distinction between smiling, or feeling a twinge of pleasure, just *when* the music takes this turn and smiling *at* the turn is the distinction between a smile or twinge that is accompanied by an underlying incident in "phenomenal consciousness" – an experience of the transition – and one that is not'.

Such a line of thought is not, on the face of it, promising. In so far as there is room for a distinction between an expressive and a merely coincidental foot tapping – between my tapping *as* the music plays and my doing so *in response to* the music – there will, on the face of it, equally be room for such a distinction in relation to the postulated intermediary: the *experience* of the music. That is to say, the very distinction that we seek to mark by appeal to an intervening act of phenomenal consciousness equally arises in connection with the relation between the manifestations (both public and private) and the act of phenomenal consciousness that is said to be the experience of change.

It is intelligible to suppose that, in a particular case, two people may display the same manifestations and yet one be experiencing the melody and the other not. That is ground for insisting that we cannot identify the experience of change with its manifestations. We need not, however, conclude that we must identify something *else* with the experience of change (and if the reasoning of the previous paragraph is sound we had better not do so on *these* grounds). For we need not assume that there is a philosophically illuminating answer to the question, 'What does the experience of change consist in?' or 'What is it to experience change?' We will only suppose that there *must* be such an answer (as opposed, for example, to the kind of answer that a German with an incomplete mastery of English will need) to questions of the 'What is …. ?' form if we are working with a fixed set of categories of 'kinds of thing that exist' into which everything must be fitted; or, as we might equally express this, with a picture of the kind of sense that the talk in question must have.

Compare this with the question: 'What does motion at an instant consist in?' The question may produce something of a mental cramp: may even tempt us to say that there is no such thing. If, from *this*, we conclude, with Zeno, that the arrow is *at rest* at any instant, we may well be working with the imagery, discussed earlier, of an instant as a time so short that there is, within it, no room for movement. But however that may be, the claim that it is not possible for there to be motion at an instant depends on a certain picture of the sense that talk of 'motion at an instant' *must* have. If we return from philosophical demands to the sense that

such talk *does*[16] have, the situation looks rather different. There *can* be motion at an instant. Achilles is running at 27 mph as he crosses the halfway point if the time at which he crosses that point lies within a period at which he is running at 27 mph.[17]

Similarly, the question, 'What is it to experience change?', as raised in a philosophical context, may acquire its appearance both of urgency and of sense from a model provided by the supposed momentary act of consciousness in which a single note is heard. We derive from that case a picture of the kind of sense that talk of 'hearing a melody' or 'seeing the movement of the rock' *must* have; and the solution to our puzzle comes when we see that the experience of something extended over time can be fitted within a modified version of that model: by, for example, postulating an act of consciousness stretched out over time with the melody (or a portion of it) as its content. But if, as I argued earlier, our model of the simple case was defective – if the experience of a single note itself has just those features that were the source of our puzzlement in relation to experience of a melody – then not only our solution, but also our problem is in trouble.

We need to confront our picture of the sense that talk of the experience of change *must* have with a consideration of the sense that talk of 'hearing a melody' or 'seeing the movement of the rock' *does* have. In my discussion of *change* at an instant I noted that talk of what happens *at* a particular time is dependent on the fact that this time is a point in an extended period during which things are happening. This is as true of the *experience* of change as it is of *change*. Thus, a description of what someone is hearing – what he is experiencing – *now* will involve reference to the melody (or note); and the melody (or note) is extended in time, having parts that lie in the past and future. But further, the sense of our claim that he is *hearing* the melody (in the sense we are concerned with in this discussion) is dependent on the fact that *he* persists through time. There is a stretch in the life of this human being in which, for example, he moves to the rhythm of the music that is being played. There is a stretch in the life of this

[16] And to the extent that Zeno is not interested in *that*, we can, perhaps, accept his conclusion 'There is no such thing as motion' – whatever *that* means – with equanimity.

[17] This is not a *necessary* condition. We have ways, greatly refined by the mathematicians, of specifying a speed at which something is travelling at an instant when that instant is *not* contained within a period during which it is moving at uniform speed. While I suspect that this fact contributes to philosophical illusions that arise here, this is not a point that I can dwell on now.

human being in which he follows the movement of the ball through the air, adjusting his position in readiness to reach out to catch it, and so on. We say of someone who behaves in these ways that he heard the rhythm, or that he saw the motion of the ball: that is what was going on in his life over that period of time. And we may say of an instant that what he was experiencing then was the rhythm of the music or the motion of the ball if that instant lay in a period of time during which he was seeing the motion or hearing the rhythm; just as we may say, of an instant, that what he was doing then was running at 20 mph if that instant lies in a period of time during which he was running at 20 mph.

We speak of what someone is now experiencing: he is, for example, hearing a melody. To see how such a description can have sense, we need not maintain that there is a sense in which the person is, at that instant, hearing notes that lie in the past or future (any more than we need say that someone who is now running the Edinburgh marathon must, in some sense, now be running through Holyrood Park). Nor need we insist that the 'now', 'the phenomenal present', is not strictly instantaneous: that the units of experience are temporally extended slices (any more than we need postulate that time itself comes in non-instantaneous chunks in order to show, in the face of Zeno, how motion is possible). To see that talk of experiencing change does not require such ideas, we need to abandon a certain picture of the relation between 'the experience of change' and its manifestations. The picture is one on which 'the seeing or hearing itself' is something distinct from – something that has a nature that may be investigated quite independently of – any of its public or private manifestations: one on which the sense of talk of what someone is seeing or hearing is quite independent of the fact that the one who sees or hears has a life, extended in time, in which her experience may have a rich variety of manifestations. But, as suggested earlier, there are good, independent, reasons for abandoning that picture.

6. 'Directly Given in Experience'

The following objection can be anticipated: 'You have focused on the *sense* of our talk: on the kind of sense that talk of "hearing a melody", or "seeing the motion of an arrow", has. But, while that is all very well, it completely bypasses the very thing that created the philosophical headaches in the first place: the fact that change and persistence are *directly given in experience*. It may be true that others are, in their characterisations of my experience, dependent on manifestations of it

in my speech and behaviour. It may be true that the ways in which talk of the experience of change features in our public discourse is closely bound up with the "manifestations" – public and private – of such experience. All of that granted, however, the experience itself is given to *me* quite independently of any of its manifestations. And that is to say: there *is* something – the character of my current phenomenal consciousness – that has its own distinctive nature: a nature that we may, as philosophers, seek to clarify quite independently of a consideration of any of its normal accompaniments in the form of foot tappings or twinges of pleasure'.

'Change and persistence are *directly given in experience*'. If asked 'What are you hearing now?' or 'What are you seeing now?', I do not need to reflect on the manifestations of my experience. Indeed, in most contexts, such self-observation – an observation of my foot tappings, or of the direction of my gaze – would be completely irrelevant to an attempt to answer the question; and if the answer I give were grounded in such observations that would show that I had failed to understand the question. We should not, however, conclude that an appropriate answer is one that is grounded in *another* form of self-observation: in my direct awareness of the current character of my phenomenal consciousness. For, in standard cases, what is called for is not attention to *myself* at all, but attention to what I am hearing or seeing: to the music being played, or to the motion of the ball. This is so whether the primary interest of the person asking the question lies in what is going on in the locality in which I happen to be or in *me*. Either way, a serious attempt to address his question does not call for *self*-reflection at all; and so we need not conclude from the fact that it does not call for reflection on the manifestations of the experience that it calls for reflection on some feature of my mental life with which I am more immediately and intimately acquainted. My ability to report, straight off, that I am seeing a stone flying (or 'the flying of a stone')[18] towards the window of that house over there is no more a reflection of a form of *self*-knowledge than is my ability to report, straight off, that there is a stone flying towards the window of that house over there.

The other learns something about *me* from my sincere attempt to describe what I am seeing or hearing: to describe, that is, what is going on around me. He learns that I am seeing the motion of the ball. In different – perhaps slightly special – contexts he may put his question to me in a way that calls for an explicitly autobiographical answer. He wants me to tell him, not what is going on around me, but

[18] It is, I suspect, a significant fact that such locutions are strained.

how I am experiencing it. It is common knowledge between us that the hand is moving, or that this melody is being played. His question is: 'can I see the movement; can I hear the melody?' There are, of course, other ways in which he might try to answer his question: for example, by noting the shifts in my gaze, the tapping of my foot, and so on. But we may (and sometimes with good reason) think of *asking* as being the most direct way of his finding out what he wants to know. Now, as in the more familiar case of the previous paragraph, there is no question here of my examining the manifestations, public or private, of my experience.[19] And so it may seem that those manifestations are simply *consequences* of the important thing – 'the experience itself' – to which I alone have direct access and which I report in the words 'I am seeing the motion of the rock'. It may seem that, as I expressed it, 'the seeing or hearing itself' is something distinct from – something that has a nature that may be investigated quite independently of – any of its public or private manifestations.

I spoke earlier of reports of what I am now experiencing as themselves being amongst the 'manifestations' of an experience of change. My point in speaking in this way was, in part, to highlight the fact that such reports do not have a privileged position in our readiness to say of someone that she experienced a particular change. Faced with a radical, and persistent, conflict between, on the one hand, a person's assurances that she hears melodies and, on the other, her inability to sing along with them and so on, we may feel we have little choice but to conclude that her assurances cannot be taken at face value – and we may do so without casting any doubt on the sincerity of those assurances.

That fact is linked with something else that I wish to highlight in speaking of 'manifestations' in the broad way that I have. Consider manifestations of these kinds: he taps his foot to the music, he follows the ball with his eyes, he feels a twinge of pleasure when the music takes this turn, or of fear when the ball swerves towards his head. We do not, I take it, suppose that such responses are an expression of the individual's intimate acquaintance with a certain feature of his current inner life: that his unhesitating tapping is to be explained in terms of a certain species of self-knowledge. But there is no more reason to suppose that his utterance 'I am seeing

[19] This is only so in the case of *present tense* self-ascription. When speaking in the past tense – of how I *experienced* something – I may appeal to its manifestations in much (though not exactly) the same way as others do when speaking of me.

the movement' should be taken to be an expression of his intimate acquaintance with a feature of his current inner life. That I can say, without investigation of any manifestations, that I hear the melody no more demonstrates that the hearing itself is something that has a nature that may be investigated quite independently of any of its public or private manifestations than does the fact that I can tap along to the music. What it demonstrates, rather, is an aspect of the way in which, given a certain form of upbringing, we may become at home in a language: may become such that my cry 'I can hear it!' reveals my state as immediately as does my involuntary start when the music takes its surprising turn.[20]

7. Consciousness in Time

I have, throughout this discussion, appealed to features of human life in time – where this includes (as it must if it is to be at all adequate) ways in which we *speak* of our experience of time: the way, for example, in which our descriptions of what is happening *now*, or what someone is experiencing *now*, are inseparably linked with the fact that this instant lies within a period during which things are happening and people are responding to them. That, it may be replied, leaves us with the task of understanding how a life with these features is possible. Clearly it is dependent on the fact that *we experience change*: beings who did not do so would never develop the ways of speaking of experience on which I have focused. To understand how our lives are grounded in such experience it must, it may be concluded, be possible to characterise this experience in ways that are uncontaminated by considerations drawn from an understanding of the world that is made possible by it.[21] We must, as we may express this, grasp how time can be present in consciousness without reference to the fact that consciousness is located in time.

[20] None of this is to deny that there are important differences between central forms of linguistic manifestation and many other forms of manifestation. There is, for example, much more ready room for a 'discovery' that I am tapping my foot than there is for a 'discovery' that I am saying 'I can hear the melody'. Such differences may, for certain purposes, justify a resistance to grouping these together under the umbrella term 'manifestations'. They do not, I think, cast doubt on what I have suggested we may learn from placing them together.

[21] In Husserl's terminology, the phenomenological investigation involves a 'bracketing' of our folk and scientific beliefs.

The conclusion simply does not follow. It is a necessary condition of an individual's having a conception of change that she has *experience* of change. It does not follow that in characterising the relevant *experience* we cannot draw on the forms of understanding of the world for which that experience is a condition of the possibility. We will only suppose that it does follow if we take it as given that there is some deep sense in which we ourselves are not part of that world; or that our experience of, for example, change provides a *justification* for the ways in which we speak and think of change: for our having a language in which we speak of change, and the experience of change, as we do. And those assumptions are not obligatory.[22]

University of Wales, Lampeter

[22] I have benefited greatly from comments on earlier drafts of this paper by a number of people. Barry Dainton and Julian Kiverstein helped me to see that in an earlier version I did not have the target of my criticisms in proper focus. I hope that the present version does better in that respect. I am also very grateful for invaluable input from Lars Hertzberg, Maureen Meehan, Soren Overgaard, Robin Le Poidevin and Angus Ross.

Quality, Thought and Consciousness

HOWARD ROBINSON

Abstract
My objective in this essay is to argue for two things. The first is that intellectual mental states – thoughts – are not physicalistically reducible, just as qualia are not reducible. The second is that thoughts and qualia are not as different as is sometimes believed, but not because – as some empiricists thought – thoughts are qualia-like by being mental images, but because qualia are universals and their apprehension is a proto-intellectual act. I shall mainly be concerned with the first of these topics.

1. Introduction

My jumping-off point for discussing the irreducibility of thought is Dennett's remark that the brain is a syntactic and not a semantic engine. Anything that is purely physical is a purely 'syntactic engine', *i.e.* it does what it does solely in virtue of its physical properties, not in virtue of semantic properties or meanings.[1] This is a truism about anything physical. It gives rise to a question for the physicalist, namely how he should cope with the psychological realization or reality of meaning and semantic properties.

There seem to be four possible responses. (1) One is Dennett's instrumentalism or interpretationalism. This is the idea that the physical syntactic engine is rendered semantic by being interpreted as an intentional system.[2] One might make an analogy with the inscriptions in a book, which are merely physical marks, but which are endowed with meaning by the interpretation that we place on them, or with the thermostat that we say turns the boiler off when it believes the temperature has reached a certain point. (2) Another

[1] As others have remarked, this is an eccentric use of 'syntactic', for syntax is hardly more of a physical feature of sentences than is semantics. The point of the metaphor (if that is what it is) is that a computing machine works because shapes fit holes, not because meanings fit anything.

[2] This idea is *passim*, at least in Dennett's early work, such as *Brainstorms* (Hassocks: Harvester, 1978) and *The Intentional Stance* (Cambridge Mass.: MIT Press, 1987). One might cite in particular 'Three Kinds of Intentional Psychology' in R. Healey (ed.), *Reduction, Time and Reality* (Cambridge: Cambridge University Press, 1981), reprinted in *The Intentional Stance*.

doi:10.1017/S1358246110000135

response is a behaviourist or functionalist realism. In this case, without essential reference to the role of an interpreter, the kinds of behaviour that the syntactic engine produces constitute meanings. One might say that, though the computer in the robot's head is a syntactic engine, the robot itself is a semantic machine, because its behaviour is intrinsically meaningful. (3) A third response is eliminativism: a denial that there are, in the end, any semantic or meaning properties. (4) Finally, and not quite a form of physicalism, is meaning-epiphenomenalism: embarrassedly accepting that meaning content, like qualia, cannot be wished away, but arguing that it carries no clout in the behaviour of people.

I shall not be considering (3) and (4), partly for reasons of space and partly because I accept what is in fact the common intuitive objection to both. When it comes to sensations, a sharp pain makes it clear both that there *is* such a thing as a sensation, and that this plays a role in my reaction. This seems to me to be even more obviously true for the content of thoughts. When I listen to what someone says in a philosophical argument, and make a reply, it seems very obvious that *what I take them to mean* plays a major role in determining what I say. One would need extraordinarily powerful reasons either to think that there was nothing that they meant, or that it had no influence on what came out of my mouth. I realize, of course, that much more could be said on this, but that is not the path I shall be following here.

I shall begin by discussing the Dennettian approach and this will naturally lead to discussion of the behaviourist/functionalist realist alternative.

2. Dennett's Instrumentalism

I shall give what I think to be the rationale for instrumentalism (or interpretationalism) and then the reason for thinking it viciously regressive.

Argument for Dennett's position:

(1) All actual intentional systems are purely physical. (Ass. of physicalism)

(2) Nothing physical is intrinsically intentional – there are no physical semantic engines. (Ass.)

Therefore,

(3) No actual intentional system is intrinsically intentional. (1, 2, HS)

(4) Actual people (*etc*.) are intentional systems in some sense. (Ass.)

Therefore,

(5) There are actual intentional systems. (4, Inst.)
(6) Actual intentional systems are not intrinsically intentional. (3, 5, MP)
(7) The only options are intrinsic intentionality and intentionality instrumentally, *i.e*. by interpretation. (Ass.)

Therefore,

(8) Actual intentional systems – people *etc*. – are intentional instrumentally, *i.e*. by interpretation. (6, 7, DS)

Reductio of Dennett's position:

(9) Something can interpret x as an intentional system only if that something has the capacity to so interpret. (Ass.)
(10) Something cannot have the capacity so to interpret solely in virtue of being itself interpreted by something else. (Ass.)

Therefore,

(11) Something cannot interpret x as an intentional system solely in virtue of its being interpreted by something else. (9, 10, HS)

Therefore,

(12) An interpreter must have the capacity to interpret intrinsically or in its own right. (7, 11, DS)
(13) Something having this capacity intrinsically is an intentional system in its own right. (Def.)

Therefore,

(14) If there are any interpreters, there are intrinsic intentional systems. (12, 13, HS)
(15) There are interpreters. (Ass.)

Therefore,

(16) There are intrinsic intentional systems. (14, 15, MP)
(17) 8 contradicts 16.

Therefore at least one of the assumptions must be false. Some are not controversial. (9) is a platitude – one cannot do something (systematically, at least, and we are considering a systematic ability) unless one

has the appropriate capacity. (7) is justified if being imputed by interpretation and having intrinsically are the exhaustive options, which seems right. (13) seems to be a correct definition. The serious options are, therefore: either to deny (1) and affirm that some intentional systems are not physical; or deny (2) and say some physical things are intrinsically intentional; or deny (10) and affirm that the capacity to interpret can itself be endowed by interpretation.

(1) is controversial, because it is the assertion of physicalism, but is the victim of the *reductio* only if (2) and (10) can be supported. As my purpose is to refute the physicalist account of thought, I need to show that (2) and (10) are true.

Denial of (10) is the assertion of Dennettian physicalism, although, as we shall see later, it can be argued that all physicalists are committed to some version of interpretationalism. Denial of (2), on its most natural interpretation, is the claim that reductive physicalism saves the realism of the mental.

3. Discussion of (10)

(10) is the anti-Dennett crux. It seems intuitively obvious: simply having a certain attitude to something cannot endow it with powers it does not otherwise possess. If an object cannot think 'in its own right' then understanding it in a certain way will not give it this ability. This intuition is, I think, sound, but diversionary tactics are possible.

It might be argued that the picture of individuals endowing others with semantic capacities, like a particular person reading a text or using a computer, is too individualistic. The interpretation is a mutual and social operation. Many social properties, including ones that endow people with powers and capacities, are endowed or imputed by what one might broadly characterize as the attitudes of others. This is the way one is endowed with legal powers. Even the power of leadership, considered as relatively brute rather than merely legal, comes from the response of others as well as from natural capacity. In some way, if Dennett is right, we must be similarly endowed with the capacity to be a semantic engine.

The natural response to this is to make a distinction between *natural powers* and *socially imputed powers*, and to claim that the power of thought belongs in the former category. This claim could be buttressed by arguing that a general power of thought is presupposed by the socially imputed powers – it is because we are intentional systems that we can construct legal systems, endow each other with powers, rights, *etc*.

The interpretationalist might try to deny that the distinction between natural and social powers is a precise one – from a physicalist perspective, after all, social powers must be natural powers, for there is no other kind, so they must just be very complex and sophisticated ones. This response is problematic for a Dennettian, however, because it appears to be at the heart of this position that semantic properties are imputed rather than real. Nevertheless, it might be argued that, though semantic properties have no place in the 'basic' physicalist ontology, they can be thought of as emergent when physical behaviour develops 'real patterns' of the right kind.

The question is what one means by 'emergent' here. The issue is how semantic facts are grounded in physical ones. Or perhaps it would be more accurate to say that it is about how semantic concepts get a grip on physical reality: are they an interpretation of that reality made for certain purposes, or do those facts emerge at a certain level of complexity, in a way analogous to that in which biological facts might be thought to emerge?

It might seem that realistic emergence is inconsistent with the interpretationalist approach, as I suggested above. But perhaps it is not that simple. One might defend a collective realism, according to which there is a society with a certain 'form of life' in which individuals really exhibit intentional states, but only in virtue of the responses that other individuals in the society have to them. The web of mutual interpretation is a real, natural phenomenon, but it is constructed by mutual interpretation; it does not emerge as a power of people taken individually.

This approach seems to me to be mere 'hand waving'. The problem concerns how more and more sophisticated patterns of behaviour come to *constitute* acts of interpretation. One needs to distinguish between:

(a) The capacity to interpret others can only *develop* or be actualized in the context of appropriate complex behaviours which we dub 'social practices'.

and

(b) The capacity to interpret others is a *logical product* of certain behaviours, such that once the patterns of behaviour reach a certain point they constitute mutual interpretation.

The former is not controversial but tells us nothing about the analysis of semantic capacities, only about a causally necessary condition for our acquiring them.

The latter fails to explain how physical behaviours transmute themselves into interpretative acts. Of course, once behaviour reaches a certain complexity, it can be *interpreted* as semantic and intentional, but this just takes us back to the start: what we want to know is how certain physical movements can simply *be what it is to interpret*.

When interpretationalism has reached this point, it is not clearly different from realist reductionism, except that the reduction base is social, rather than individual, behaviour. We will return to discussing this in section 5.

4. Realism and 'Real Patterns'

Dennett came to dislike the label 'instrumentalist' and to declare that he was a 'moderate realist'. He expressed this by saying that the intentional states and systems whose value he had originally described as *instrumental* were, or were grounded on, *real patterns* in the physical world. How is this position related to the realism to which I have driven the interpretationalist in the previous section?

The current discussion concerns whether instrumentalism/interpretationalism involves a vicious regress. The regress reputedly consists in the fact that this theory presupposes an intrinsically intentional system to carry out the interpreting required by both; it cannot, therefore, constitute an explanation of what it is for something to be an intentional system or a semantic engine. Dennett's account of real patterns does not seem to me to touch this accusation. In order to meet the challenge he would have to show that a pattern, on its own and without the aid of some act of interpretation, constituted an intentional state. One can make a distinction between the view that patterns are *per se real* and that they are *grounded*. On the latter account, a pattern is a kind of *Gestalt*, because it is a matter of a certain structure being seen as a whole in a certain way. The figure created by a continuous line moving equidistant from a central point just *is* a circle. A series of dots placed on the same outline as the circle will also be seen as forming a circle, but they are just dots in certain positions: they could be seen – if they were seen as forming anything other than a collection of dots in certain places – as forming a polygon. The pure circle is not a pattern, it is a self-sufficient shape. The dots form a pattern which requires a mind – an interpreter – to complete it, and it could be completed in more than one way, though one particular way may be the easiest or most natural. William Seager, in an illuminating discussion of

Dennett on real patterns almost gets this right, but not quite and this lets Dennett off the hook. Seager explains the status of patterns as follows:

> Inhabiting a curious zone midway between, as it were, objectivity and subjectivity, patterns are *there* to be seen, but have *no function* if they are not seen. By the former I mean that patterns are not just in the eye of the beholder; they are really in the world and provide us with an indispensable and powerful explanatory and predictive grip on the world. By the latter I mean that the *only* role they have in the world is to help organize the experience of those conscious beings who invent them and then think in terms of them.[3]

It generally looks as if Seager is saying that patterns are really there, but are physically epiphenomenal, because all the causal clout comes from

> …the fundamental features of the world [that is, its most 'minute parts'] that organize the world into all the patterns it exemplifies, and they do all this by themselves, with no help from 'top-down' causation.[4]

Seager seems to think that patterns are not wholly epiphenomenal because they are picked out by and hence influence minds, and that, for this reason, 'Mind cannot be "just another" pattern'.[5]

Dennett has a twofold reply. First, he rejects the view that patterns are physically idle: ('All those simpler, thermostat-like minds are responsive to patterns…'[6]). This, in a sense, is a verbal dispute, for Dennett is not denying that the world is 'closed under physics' and that, therefore, higher order entities *add nothing* to the causal clout of the minute parts. Nevertheless, Dennett is right that this latter fact does not seem to make it wrong to attribute causal force to non-fundamental entities: it is still the stone that broke the window, even if this supervenes on the action of the atoms. Second, and more crucial, he rejects Seager's main conclusion:

[3] W. Seager, 'Real Patterns and Surface Metaphysics' in D. Ross, A. Brook and D. Thompson, *Dennett's Philosophy: A Comprehensive Assessment* (Cambridge, Mass.: MIT Press, 2000), 95–129: 117.

[4] *Ibid*.

[5] *Ibid*., 121.

[6] D. Dennett, 'With a Little Help from My Friends' in D. Ross, A. Brook and D. Thompson (eds), *Dennett's Philosophy: A Comprehensive Assessment, op. cit*., 327–88: 355.

> In Seager's opinion, 'Mind cannot be "just another" pattern' Why not? Perhaps I have missed his point.[7]

Dennett has missed the point but only because Seager has not stated it quite correctly, or not clearly so. If patterns are real and are of the same ontological status as higher order, non-fundamental objects in general, and if it is appropriate to ascribe causal roles to such non-fundamental things, even though these supervene on the atomic, then why should not the mind be efficacious, as the stone is, and just a pattern? Seager's mistake is to characterize patterns as real but inefficacious, except upon the mind.

In fact he is ambiguous about the reality of patterns. In one of the passages quoted above he talks of patterns as *invented by* conscious beings. Seager needs to make explicit what he perhaps intends, namely a distinction like that I make above between the *groundedness* of patterns in reality, together with the need for mental activity to reify them on the basis of those grounds. This explains why the mind cannot be just a pattern: it is presupposed by patterns as their co-inventor, together with the grounding. If the mind itself were just a pattern, then there would be the kind of regress with which we started our discussion, for it would not be reified unless it were *seen as* a pattern, and so on.

5. Social Realism

So Dennett's doctrine of real patterns does not enable him to escape from the regress. But we ended section 3 with a form of social realism-cum-interpretationalism still in play, and must return to the discussion of that theory. In fact, the problems that we found for the real patterns theory also apply here. We saw that Seager, in his rejection of 'top-down' causation was implicitly classifying all non-foundational entities as similar to patterns: that is, he did not demote patterns because they were patterns *per se*, but because they were higher order entities and higher order entities had no independent causal clout, because all such clout is 'bottom up', not 'top down'. This assimilation of all non-fundamental entities to patterns is, I believe, essentially correct. I have elsewhere presented an account which can briefly be summarized as follows.[8]

[7] *Ibid*.

[8] H. Robinson, 'Dualism' in *The Blackwell Guide to the Philosophy of Mind*, edited by S. Stitch and T. A. Warfield (Oxford: Blackwell, 2009), 85–101; H. Robinson, 'Reductionism, Supervenience and Emergence' in

There are two forms of strictly 'bottom up' reductionism. One is the 'translation' reductionism of Carnap[9] and other logical positivists, according to which all true statements in the special sciences and commonsense ontology can be translated into statements about fundamental physics. The other is the 'nomological reductionism' associated with Nagel's classical account, according to which higher order properties and laws are type identical with something in physics.[10] In both these cases the conceptual or explanatory content of higher order descriptions adds nothing to what can, in principle, be acquired from a proper account in terms of physics. Unfortunately, neither of these forms of reduction actually applies to the relationship between physics and most, if not all, higher order descriptions. Even if and where the world is 'closed under physics', the relation between that fundamental physical base and the rest is only a form of *a priori* sufficiency of the base; there is not also the necessity of the base that either translation or nomic reduction requires. By '*a priori* sufficiency of the base' I mean the following. Given what is happening at the fundamental level, then what is happening at the higher order levels follows necessarily. For example, though there is no nomological reductive account of 'hurricane', given that the atoms are behaving in a certain way, then necessarily there is a hurricane: there is no possible world atomically just like ours at the time of Katrina in which there was not a hurricane. This is so even though the conceptual frameworks of the higher explanations, such as meteorology, cytology, *etc*. 'float free' of the conceptual framework of physics. I argued that this shows that the special sciences are best understood as different perspectives on the physical base, usually with certain interests in mind. They are essentially in the same category as patterns, because, though the concepts they involve are well grounded by the basic physical reality, they do not reflect any reality additional to the fundamental physical base, except the interests and other perspectives of the humans who employ them. These perspectives do not differ significantly from modes of interpretation of the patterns available at the lower level. In other words, given the failure of tough minded forms of reductionism, the relation between the base and other levels of explanation is 'top

---

The Routledge Companion to Metaphysics, edited by P. Simons and R. Le Poidevin (London: Routledge, 2009), 527–36.

[9] R. Carnap, *The Unity of Science* (London: Kegan Paul, 1934).

[10] E. Nagel, *The Structure of Science* (London: Routledge and Kegan Paul, 1961).

down', and this is a form of interpretationalism, which presupposes a mind picking out the *fundamenta* that make the higher order explanations possible.

If the above argument is correct, all physicalists are interpretationalists, not just about mental states, but about all, or, at least, most of the special sciences (that is, those not reducible in one of the strong senses). Premise (2) of the original argument – the denial of straight realism about the mental, from a physicalist perspective – is correct. But I have already shown, in my defence of (10), that interpretationalism leads to a vicious regress. Neither physicalist realism nor interpretationalism is able to accommodate thought. The other options open to a naturalist that I mentioned at the outset were eliminativism and meaning epiphenomenalism. Given that neither of these is acceptable, one is forced to a realist, interactionist, and dualist theory of thought.

6. Why It Is impossible to Combine a Non-reductive Account of Qualia with a Reductive Account of Conceptual Activity

There is another strategy for showing the irreducibility of intellect, namely to show that it follows from the irreducibility of phenomenal content. By contrast, it is not unusual for philosophers to combine a non-reductionist acceptance of qualia with a functionalist or behaviourist account of thought. The general view of those who do this is that the reductive approach works for everything except 'raw feels'. Examples of versions of this approach are Ayer, Jackson, Chalmers and, I think, Russell.[11] It seems to me that this combination is impossible; if you accept the irreducibility of phenomenal content, you must accept the irreducibility of at least certain basic intellectual acts, namely those involved in recognition. My argument for this is, in outline, as follows:

(1) The view I am attacking combines (a) the irreducibility of qualia, with (b) a behavioural/functional account of conceptual activity, and, hence, of recognition.

[11] A. J. Ayer, *The Origins of Pragmatism* (London: Macmillan, 1968); F. Jackson, 'Epiphenomenal Qualia', *Philosophical Quarterly* **32** (1982), 127–36; D. Chalmers, *The Conscious Mind* (Oxford: Oxford University Press, 1996); B. Russell, *The Analysis of Matter* (London: Allen and Unwin, 1927).

(2) This requires that the having of the qualia and the act of recognizing them be separate acts or events, and, consequently, they are only causally and contingently connected.

(3) If (2) is true, then the nature of the qualia could vary without the content of the act of recognition varying, because they are only contingently connected.

(4) This would lead to 'beetle in the box' redundancy. That is, if what it seemed to the subject he was recognizing could float free of the nature of the qualia, then the qualia would be redundant.

Therefore,

(5) Qualia and the associated act of recognition are not contingently and causally connected.

Therefore,

(6) Qualia must be ostensively internal to recognitional judgements. That is, as Russell maintained, in a recognitional judgement such as '*that* is red' the quale being demonstrated is part of the content of the judgement.

The transition from (1) to (2) can be illustrated as follows. Suppose the subject to be veridically perceiving a red object. Using an arrow to represent causal connection, the process would be as follows:

Red object → red quale → red-appropriate functional/behavioural response.

Suppose instead the following occurred:

Red object → green quale → red-appropriate functional/behavioural response.

What are we to suppose it would be like for the subject in this latter case? *Ex hypothesi*, noticing that the quale is green would involve recognition, but all the recognitional and conceptual responses are of the red-appropriate kind, so the subject cannot notice the nature of the quale: it is not to him as if it were green. Suppose the following occurred:

Red object → no quale → red-appropriate response.

For the same reason, the subject cannot notice that there is no content to the experience. So the causal, naturalistic account of recognition leaves the hypothesization of qualia redundant. In other words, a property dualist account of experience or phenomenal content must

incorporate some basic form of recognitional apprehension of that content, otherwise its whole purpose, which is to be the essence of 'what-it-is-like' for the subject, will be evacuated. To use Wittgenstein's famous image, it would not matter whether or not there was a beetle in the box.

7. Understanding Meaning: Proliferation and 'Magic'?

If we are going to accept that conceptual activity is irreducible, just as is phenomenal consciousness, then are we not allowing a proliferation of distinct irreducible entities, and/or capacities? The thought behind this worry is that phenomenal consciousness, or 'raw feeling', is something entirely different from thought: thought essentially involves concepts which are, in some sense, universals, but phenomenal consciousness concerns only a strange, private kind of particular.

Whilst I do not want to deny the importance of the difference between thought and sensory experience, I think that the respect in which there is continuity between them can be missed. The classical empiricists tended to assimilate them by having an imagistic theory of thought, thus moving in the direction of reducing thought to sensation. My suggestion indicates the opposite direction. It is important that phenomenal contents are qualitative in nature; that is, they are universals, though not abstract objects. I have argued above that, in the case of human experience, at least, it is not possible to fix a clear divide between having the experience and the minimally conceptual act of recognition. This might seem to imply that in experience we apply concepts to our raw feels, and it is this process of forming concepts that seems problematic and wholly different from simply having the experience. But once one recognizes that phenomenal contents are qualitative and, hence, essentially universal in nature, one can see that merely apprehending them is a proto-intellectual act.

I think that the divide between sensory and intellectual acts can be drawn in either of two ways, one of which I shall dub 'nominalist' and the other 'Aristotelian realist'. According to the nominalist, any form of conceptualization has to be constructed, because there is nothing universal in nature, and, somehow, the mind has to construct generality for its own purposes. In my opinion, there is no way in which this could be done, but the idea that this must be how conceptualization works forces a division between the absolute particularity of experience and the more or less linguistic sophistication of anything

conceptual. For the realist, on the other hand, all properties and qualities are at least immanently universal: a mind does not so much need to construct a universal to fit a quality it senses, as to apprehend a universal that is already present. All discernment of similarity and dissimilarity between phenomenal contents – without which there seems hardly to be experience at all – is, if the realist is right, already a grasping of universals. On this theory, the great divide comes, not between phenomenal contents as pure particulars and concepts as constructed universals, but between the ability to apprehend universals only when they occur instantiated sensorily in experience (or quasi sensorily in images) and the ability to grasp them *in absentia*, as abstract objects in thought. This latter seems to require syntactic structure and, hence, language.

There is, of course, much more that needs saying about this approach, but it does help to undermine the idea that taking a grasp on meanings as irreducible is some form of 'magic'. Perhaps surprisingly, both Putnam and Kripke, who are anti-physicalist, decry the idea that the semantic properties of mental states are irreducible as 'magic'.[12] Kripke does this in the course of his discussion of Wittgenstein, and his refusal to countenance our grasp on meanings (or intensions or universals) leads him to an extreme conventionalism. The argument is as follows:

(1) Because we cannot grasp open-ended, potentially infinite extensions in an extensional manner – that is, by running through the potentially infinite extension individual by individual - either we can grasp intensions *per se* or concepts are constructed conventionally/nominalistically in a Goodmanian way: that is, for every new case it is a decision whether to include it within a certain kind.

(2) Grasping intensions is unacceptable because it is 'magical' and 'superstitious'.

Therefore,

(3) Concepts are constructed in a Goodmanian way.

If this is correct, there are no natural kinds or similarities as such, only ones made up by stipulation. If you think (rightly) that this is absurd, then you will work a *modus tollens* on the above *modus ponens*:

[12] S. Kripke, *Wittgenstein and Rule-Following* (Oxford: Clarendon Press, 1976); H. Putnam, 'Brains in a Vat' in his *Reason, Truth and History* (Cambridge: Cambridge University Press, 1981), 1–21.

(4) Goodmanian nominalism/conventionalism destroys the whole rationale of concepts and thought.

Therefore,

(5) We must grasp intentions as such.

The thought that this is 'superstition' is connected with the idea that experience only puts us in touch with particulars, of which we would have to survey an infinite number to collect them under a genuine universal. But once one realizes that experience is of essentially universal entities, then no anti-reductionist about experience is going to regard the ability of the conscious mind to grasp intensions as an extra 'superstition': consciousness and the grasp of things as kinds are hardly distinguishable.

8. Conclusion

There is no acceptable physicalist account of our intellectual abilities. This is so for two reasons. First, all physicalist accounts are forms of interpretationalism and the intuition that these accounts lead to a vicious regress is correct. Second, there is no physicalist account of phenomenal content and the irreducibility of the intellectual follows from this, because of the role of recognition in the 'what-it-is-like' of phenomenal content. Furthermore, the divide between the phenomenal and the intellectual is not as absolute as sometimes thought, because in both cases it involves apprehending universals – or, in Aristotle's phrase, the reception of form without matter.[13]

Central European University, Budapest

[13] I am grateful to the participants at the *Metaphysics of Consciousness* conference in Edinburgh in July 2009 and to Anita Avramides and Adrian Moore who commented on a later version of the paper.

Concessionary Dualism and Physicalism

WILLIAM SEAGER

Abstract
Modern physicalists frequently offer the generous concession that although dualism is false, it is not a metaphysical impossibility. And it appears that the proper formulation of physicalism allows for this concessionary position. It would be expected that dualists also could accept that while physicalism is false, it too is a metaphysical possibility. I will argue that a careful analysis of physicalism and dualism shows that in fact these concessionary positions cannot be maintained. In particular, the nature of the metaphysical determination relation which holds between matter and mind on both physicalist and dualist views precludes either from allowing that the other is a metaphysical possibility.

1. What is Physicalism?

The doctrine of physicalism can be roughly spelled out simply as the claim that the physical state of the world determines the *total* state of the world. However, since there are many forms of determination, a somewhat more precise characterization is needed.

One obvious problem with the simple formulation is that the traditional doctrine of epiphenomenalism holds that the mental is determined by the physical. However, the orthodox view, which seems obviously correct, is that physicalists would and should balk at the claim that epiphenomenalism is a form of physicalism.

The philosophical zombie thought experiment vividly reveals exactly why epiphenomenalism is not a version of physicalism. According to traditional epiphenomenalism the determination relation in question is causation, and causal relations do not hold with full necessity. They hold with at most nomological necessity. Thus, if epiphenomenalism is true, there is a possible world, w, in which the causal laws are different in such a way that the physical states which in the actual world cause mental states either cause no mental states at all in w (the zombie option) or cause aberrant mental states (the inverted spectrum option). Why should the physicalist care about this 'mere possibility'? Because it shows that the mental can vary independently of the physical and there can be no

doi:10.1017/S1358246110000147

better demonstration of ontological distinctness than independent variation.

Therefore, physicalism requires that the sort of determination at issue must exhibit *maximum* modal force; it must be absolutely impossible for the mental to vary without attendant, determining physical variation (let us label this relation 'logical determination'). I think the best way to state physicalism which meets this constraint is in terms of what are called minimal physical duplicates (MPDs) of possible worlds, an approach pioneered by David Lewis.[1] An MPD of the actual world is a possible world which is physically exactly similar to the actual world and contains nothing else. In other words, an MPD of the actual world is configured in just the way that physicalists assert the actual world is configured. Physicalism is true of world w, then, just in case w is its own MPD, or, more formally (letting 'Px' stand for 'physicalism is true at world x' and 'Mxy' for 'x is an MPD of y'),

$$\mathbf{MPD-P}. \quad (\forall w)(Pw \equiv (\forall v)(Mvw \rightarrow v = w)).$$

Obviously, the identity condition rules out the existence of *any* possible world which exhibits independent variation of non-physical features relative to the physical. So we have the desired maximum modal force. At the same time, a very nice feature of this definition is that it permits physicalism to be a contingent truth about the actual world without weakening the modal force of the determination relation. That is, even if physicalism is true of the actual world there are still other possible worlds where it is false. A world where Cartesian dualism was true would be a world where physicalism is false, but it may for all that be a genuinely possible world, and it would seem not to threaten the truth of physicalism.

The definition nonetheless does have some odd consequences. It utterly fails to even hint at the nature of the link between the fundamental physical features of the world and the higher level structures which depend on them.

One attractive way to avoid this consequence is to embrace an additional constraint on the definition of physicalism: the in principle *a priori* deducibility of the mental description (or whatever high level description is at issue) from the purely physical description, given

[1] D. Lewis 'New work for a theory of universals', *Australasian Journal of Philosophy* **61** (1983), 343–77. See also F. Jackson, *From Metaphysics to Ethics* (Oxford: Clarendon, 1998) and D. Chalmers and F. Jackson, 'Conceptual Analysis and Reductive Explanation', *Philosophical Review* **110** (2001), 315–61.

possession of the relevant high level concepts. Such an approach, though not exactly the one I outline here, is endorsed by both David Chalmers and Frank Jackson.[2] Very roughly, the idea is that if physicalism is true then there is a sense in which all truths in world w should be deducible from the purely physical description of w. For example, consider the set of chemical truths. These are not explicitly represented in the purely physical description of the world but – it is very plausible to believe – they are logically determined by the purely physical truths. What is more, it seems also to be true that given the purely physical description *plus* a grasp of the concepts of chemistry, the chemical description of the world could be deduced.

We might label the additional condition of *a priori* deducibility 'logical transparency' (the resulting characterization of physicalism would simply be that it is logically transparent that the defining condition of MPD-P holds). The motivating advantage of logical transparency in the definition of physicalism is that it provides the requisite link between fundamental and higher level features.

Logical transparency can also be regarded as a benchmark of the explanatory adequacy of physicalism. If physicalism is true we expect that, in principle, there *should* be an account of how or why higher level concepts apply to the world. We can hope that it is accessible to us even though the ideality condition means there is no guarantee.

Another problem with MPD-P is that it threatens to make physicalism come out trivially true under certain conditions in which physicalism is, intuitively, false. Consider a universe, call it Diffuse, exactly like ours with respect to physical law but in which the initial conditions of the big bang left it in a state where the originally created elements ended up so smoothly and sparsely distributed that stars and galaxies could never form. It is easy to believe that any MPD of Diffuse is completely identical to Diffuse – there is just nothing going on in that world except basic physical processes. So physicalism comes out true in Diffuse. But I think this is worrisome. This imaginary world seems as capable of possessing the intrinsic 'power' needed to generate biology and mentality as does the actual world.

Now, suppose that physicalism is false of the actual world. Then I think it should also be false of Diffuse as well. Could champions of physicalism make the doctrine true simply by wiping out all

[2] D. Chalmers and F. Jackson, 'Conceptual Analysis and Reductive Explanation', *op. cit.*

sentience? One could maintain that an MPD of Diffuse will not be identical to Diffuse since it will lack some extra non-physical laws of nature, the laws of emergence if you will, that would kick into action if and when the right kind of complex physical structure should appear (which as a matter of fact will never happen in Diffuse). But then one might worry about the status of fundamental laws (that is, laws which are not derivable from other laws) which are never instantiated in a world. And one might have Humean scruples in favor of the idea that laws supervene on regularities, and the requisite regularities will never appear in Diffuse.

There is a simpler remedy which avoids these metaphysical conundrums. It also has the advantage of explicitly recognizing that physicalism is best regarded as a doctrine about whether certain specific target phenomena are logically determined by the physical. Physicalism in general is the claim that all phenomena are logically determined by the physical, but more restricted forms are possible, and the general form is nothing more than the logical sum of all the restricted forms. The present position of strength which physicalism enjoys in philosophy stems from the relentless assimilation of ever more domains of phenomena under the purview of physicalism.

Let us then take physicalism to be the domain relative claim that for any possible world and domain of facts, d, there are some d-truths which hold at that world and all its d-truths are true at any MPD of that world, or, more formally (using 'Pdw' for 'physicalism with respect to domain d is true at w', 'p_d' for d-propositions and 'Txy' for 'x is true at world y'):

$$\textbf{MPD} - \textbf{PD} \quad (\forall w)(\forall d)(Pdw \equiv (\exists p_d)(Tp_dw) \wedge (\forall v)$$
$$(Mvw \rightarrow (\forall p_d)(Tp_dw \rightarrow Tp_dv)))$$

Thus, physicalism about the mind is only going to have a chance to be true of worlds where there is mentality. Diffuse is not a world where physicalism about the mental is true (albeit for the trivial reason that there are no minds in Diffuse). A more sophisticated definition could hold that physicalism about a domain is neither true nor false in a world where that domain is not instantiated. MPD-PD has the advantage that physicalism could turn out to be true of chemistry but false of psychology or, more specifically, true of intentionality but false of phenomenality. General, or total, physicalism is the doctrine that physicalism is true for all instantiated domains. In this case, MPD-PD reduces to MPD-P given – what seems obvious – that any two possible worlds that agree on all truths are the same world.

2. Zombies and Physicalism

The zombie argument against physicalism about phenomenal consciousness is crystal clear. It purports to show that consciousness can vary independently of the physical. There could be no better argument for ontological distinctness.

The relation between physicalism and zombies is more interesting than this, however. They are in much stronger opposition than the above argument reveals, because if zombies are possible then physicalism is not only false, it is *necessarily* false. Here is the argument for that:

(1) $$\Diamond Z \rightarrow \neg P.$$

Here, Z = there are zombies, P = physicalism is true. To say that there are zombies is, of course, shorthand for saying there are creatures which are physically identical to targeted conscious creatures in the world of evaluation but which lack consciousness. To say that physicalism is true is shorthand for saying that physicalism about consciousness is true.

(1) follows from the argument presented above. That argument was entirely *a priori*, so we can infer that

(2) $$\Box(\Diamond Z \rightarrow \neg P).$$

One might regard this move as somewhat swift. It is not after all *a priori* true that consciousness exists. But an argument for (2) is easily provided. (2) is false if there is a possible world where zombies are possible but physicalism about consciousness is true. As noted above, physicalism about consciousness is true at a world only if consciousness exists in that world. So we need only consider a world in which consciousness is instantiated, physicalism is true and zombies are possible in that world. Obviously, such a situation is inconsistent. (2) is vindicated.

From the fact that necessity distributes over implication, we deduce from (2)

(3) $$\Box\Diamond Z \rightarrow \Box\neg P.$$

It is evident (from the appropriate modal logic for metaphysical possibility, which is S5-like) that if zombies are possible then it is necessary that they are possible, that is, $\Diamond Z \rightarrow \Box\Diamond Z$, from which we can infer the desired

(4) $$\Diamond Z \rightarrow \Box\neg P.$$

What should we make of this? First, if zombies are so much as possible then there are no possible worlds where physicalism about consciousness is true.

Therefore anyone who thinks that zombies are possible is going to 'know' that physicalism *cannot* be true. The argument is not very complex, so quite minimal rational reflection on the part of someone who regards zombies as possible reveals that the truth of physicalism is inconceivable. It is inconceivable in the same sense that the idea that there is a greatest prime number is inconceivable – rational reflection shows that the concept of prime number is inconsistent with there being a largest prime number. So, it is genuinely inconceivable that physicalism is true given a commitment to the possibility of zombies.

3. Stalemate?

By applying *modus tollens* and some simple modal-negation manipulation we can derive the *physicalist's* version of (4):

(5) $$\Diamond P \rightarrow \Box \neg Z.$$

This shows that physicalism and the possibility of zombies are 'modal contraries', and the mere possibility of one entails the logical impossibility and hence inconceivability of the other. I think this does reflect the kind of 'rock bottom' level of disagreement between defenders of physicalism and those who defend some kind of dualism. Perhaps this is an irresolvable clash of intuitions.

But is the dialectical situation really one of intellectual stalemate here? I think this is unlikely to be the case and that the dualists have a distinct advantage. The argument would stall completely if the intuition in favor of the possibility of zombies and in favor of physicalism were equally strong. This does not seem to me to be the case, however. There is no basic intuition in favor of physicalism. Its strength springs from the history of success which science has enjoyed over the last 400 years or so, as more and more of the natural world has been linked to the entities and processes described by fundamental physics. Although this linkage is undeniably supportive of physicalism, it has always been incompletely specified and seems sure to remain so, if for no other reason than the staggering immensity of the details of the determination relation which holds between the fundamental microphysical world and the macrophysical world of the 'manifest image'. And it is not simply details that stand in our way; we have at present no good understanding at

all of how the observed world can emerge from the underlying quantum reality.

On the other side of the coin, the possibility of zombies is a clear extension of age old worries about the highly significant gulf between what we can observe of people and their mental states. As Shakespeare expressed it in *Macbeth*: 'There is no art to find the mind's construction in the face'.[3] The idea of the mindless automaton which counterfeits human action is also very familiar, as is the notion of a trance state which renders someone unconscious but capable of action. The philosophical problem of other minds is yet another familiar difficulty, which can be explained to non-philosophers in just a few minutes, that reinforces the apparent possibility of zombies.

The zombie conception goes much further than these considerations of course, but it is worth remembering that this thought experiment is simply the extreme end of a continuum, any point of which will refute physicalism. Instead of a zombie, we can imagine a possible world physically identical to this one but in which the precise hues of experienced colors are ever so slightly different than in the actual world. This is, on the supposition of physicalism, absolutely impossible but it is hard to see exactly what it is about the physical structures which so perfectly and definitely constrains all the details of the mental with the inexorable force of logical necessity.

It may be that future science will reveal how the existence and precise qualitative features of subjective experience are logically determined by physical states and purely physical laws. It was, after all, ignorance of the laws of physics that permitted philosophers from J. S. Mill to C. Lloyd Morgan and C. D. Broad[4] to deny that chemical properties were necessitated by underlying physical structure. But since ignorance (leavened with rational reflection) is a legitimate ground for at least tentative judgments of possibility, it seems to me that the hypothesis that zombies are logically possible, despite its philosophically extreme nature, is less epistemically extreme than the physicalist hypothesis. The basis of the former is a kind of (putative) insight into the nature of consciousness, the latter is based upon an inductive inference from an incomplete and possibly irrelevant database.

[3] Act 1, Sc. 4.

[4] J. S. Mill, *A System of Logic, The Collected Works of John Stuart Mill* (Toronto: University of Toronto Press, 1963), Vols. 7–8; C. Morgan, *Emergent Evolution* (London: Williams and Norgate, 1923); C. D. Broad, *Mind and Its Place in Nature* (London: Routledge and Kegan Paul, 1925).

4. Concessionary Physicalism and Dualism

Perhaps reflections such as the above are what prompt the concessionary spirit in physicalists, who allow that while physicalism is true, dualism is possible.[5] Recall that the definition of physicalism allows that there could be possible worlds in which Cartesian dualism provides the true account of the mind-body relation. Such worlds are logically compatible with the idea that the physical logically determines the mental so long as none of those worlds are physically identical to the actual world.

It would be natural to think that there should be an argumentative symmetry here. That is, it might be expected that just as the physicalist will concede that there are possible worlds where physicalism fails, so too the dualist would also allow that there are worlds where physicalism is true. If so, then following the argument given above it turns out that according even to the dualist zombies are impossible.

That the possibility of zombies does not logically follow from the acceptance of dualism is quite interesting. But while this may be strictly correct, it seems very uncomfortable to hold to both the possibility of dualism and the impossibility of zombies. Consider a possible variant of a standard Cartesian world, a world that is physically just like Diffuse but in which Cartesian immaterial minds exist, though they are not anchored to any physical entity (perhaps they are angels). Physicalism is clearly false at such a world because, while mentality is instantiated in this world, its MPDs will lack all mental features. There are no zombies in an MPD of Cartesian-Diffuse, but if *all* minds were Cartesian minds the path toward the possibility of zombies is not hard to make out.

Thus, the concessionary dualist who wishes to allow for the possibility of physicalism is going to have to say that while some minds are – or may be – Cartesian, others are physical. Unfortunately, it is not very clear what a dualist could mean by this.

In line with our understanding of physicalism, dualism should be conceived as a claim about the actual world: the mental features of the actual world are radically non-physical; physicalism is false of the actual world. But, I will argue, on this reading, any dualist should be strongly inclined to think that zombies are possible.

For consider that according to the concessionary dualist the possibility of zombies is prevented only by the bizarre modal fact that it is

[5] See for example F. Jackson's 'Finding the Mind in the Natural World' in *The Nature of Consciousness*, edited by Ned Block, Owen Flanagan and Güven Güzeldere (Cambridge, MA: MIT Press, 1997), 483–92.

somehow impossible for a putative MPD of the actual world to lack mental features but retain a physical identity with the actual world. But according to the dictates of MPD physicalism, that would mean that physicalism would be *true* of the actual world, not false. Dualism would not stand in opposition to physicalism in its claims about the way things actually are.

The problem is with the interpretation of the Cartesian idea that without input from the mind, physical stuff could not (and this must be a *logically* 'could not') act as it does in the actual, enminded world. Such a modal constraint is very hard to grasp. In any case, this is an incredibly thin reed to hang a doctrine upon.

In the first place, it is hard to see how the dualist could avoid the possibility of 'partial zombies'. These are creatures physically identical to a Cartesian enminded creature during a period of time in which there is no interaction between body and mind (perhaps during a period of purely intellectual reflection). Given the appropriate construction, one could define a suitable partial MPD: a short lived possible world physically identical to the actual world during the period of zero interaction but lacking all Cartesian minds. Such a world would serve to refute physicalism via the possibility of zombies. The dualist can disallow such a world only if it is logically impossible for there to be a period of time during which mind and body fail to interact.

Furthermore, a purely physical world in which there was indeterminacy (such as quantum mechanics suggests obtains in the actual world) would seem to be able to present a purely physical appearance in complete accord with that of the actual world, even if Cartesian minds are at work there. Wherever in the actual world a mind stepped in to effect a change in physical processes, let a simple jolt of indeterminacy direct the material process in the same way, though without the aid of any mental cause.[6] Arguably, such a world is an MPD of the actual world, but one without any mental features. In such a world, there are zombies.

[6] Karl Popper and John Eccles suggested in *The Self and Its Brain* (New York: Springer International, 1977) that the mind could exploit quantum indeterminacy to operate in the physical world, neural synapses being the 'seat of the soul', with no apparent violation of physical law. It follows that there could be an MPD of an Eccles-Popper world that lacked mentality; that is, that contained zombies. Of course, under the right choice of measurement sequences the Eccles-Popper world would look extremely unlikely relative to the calculated quantum probabilities of synaptic transmission, but extremely unlikely is a long, long way from impossible.

Perhaps the concessionary dualist could object that in the actual world there is no indeterminacy. Despite appearances the physical laws are all completely deterministic (but are broken by mental intervention upon occasion), and so our scenario fails to yield an MPD of the actual world. This is the same thin logical reed. It would ultimately entail that according to the concessionary dualist it would somehow be logically impossible to construct an MPD of the actual world without including non-physical mental features in it. The postulation of such a bizarre, brute modal fact does not seem to count exactly in favor of dualism.

One potential way for the dualist to support this bizarre necessity is to endorse a causal essentialist account of property identity. The dualist can then explain why there are no MPDs of the actual world which lack mental properties without giving up dualism. For it is then open for the dualist to claim that it is strictly necessary that any MPD of the actual world will instantiate mental properties, since it will be an essential characteristic of the physical features of the world (when properly arranged) that they causally generate mental features.[7] There are many difficulties with this response. One is that even if we endorsed causal essentialism for properties it would not follow that *all* causal powers of a property are essential to it. It might be that the essence of physical properties lies entirely in their power to interact with purely physical aspects of the world (even if they also possess contingent powers to generate non-physical features). A more serious difficulty stems from the option available to the physicalist to refine the definition of minimal physical duplication, restricting the causal powers of physical properties to physical effects. This seems reasonable given that physical causation resolves itself into the fundamental physical features of the world (see section 5 below). If it turns out that mentality will not arise in a world whose physical properties have been thus restricted then physicalism has been falsified even if there is a sense in which the restriction means that we have altered the identity of the physical properties which inhabit the actual world. The dualist must at this point simply insist that this restriction is somehow impossible – that it is inherent in matter that it causes non-physical effects.[8] This does seem

[7] This idea is deployed by Brian Garrett in the context of an anti-zombie argument. See his 'Causal Essentialism versus the Zombie Worlds', *Canadian Journal of Philosophy* (forthcoming).

[8] Such a dualist would be endorsing a form of what is often called Russellian Monism, but not one which could be allowed to count as a form of physicalism. See D. Chalmers, *The Character of Consciousness*

extremely implausible and thus it seems to me that a dualist ought to embrace the possibility of zombies, and hence deny the possibility – and hence the conceivability – of physicalism.

Another interesting problem then looms. The physicalist made much of the magnanimous concession to the dualist that there might be Cartesian worlds. But it seems just a small step from those possible worlds to the possibility of zombies. Now it is the *physicalist* concession which can stand only if the same thin logical reed is put in place. Cartesian worlds are possible only if it is impossible that such worlds have MPDs which lack mentality. But if Cartesian worlds are possible they will – barring bizarre brute necessities – have such MPDs. And the Cartesian world which is physically identical to the actual would will have an MPD which is an MPD of the actual world, thus refuting physicalism via the reinstatement of the possibility of zombies.

The maximum physicalist concession perhaps should be along the following lines. There are worlds just like the actual world but which include non-physical minds. Physicalism is false at such worlds since their MPDs lack the immaterial minds. But these worlds retain the mental features necessitated by the physical processes going on in them. There are no worlds in which the mental states 'associated' with physical bodies all belong to Cartesian minds which are linked to those bodies according to Descartes' picture of the substantial union. Somehow, these latter are not permissible dualist worlds.

The maximum dualist concession is … what? If dualism requires belief in the possibility of zombies, then the dualist is severely constrained. So the key issue remains whether dualism entails the possibility of zombies. As we saw above, on the face of it, strictly speaking there is perhaps no such entailment. But a closer look suggested otherwise. This because of the general nature of the mind-body relation that dualism must posit. Let us have a further look at this relation.

The general issue is the familiar one of property dependence or supervenience. A superficial survey presents a number of distinct sorts of dependence that are worth distinguishing.

(Oxford: Oxford University Press, forthcoming) for a discussion of the role of Russellian Monism in the general conceivability argument against physicalism (a version of the argument is available online at http://consc.net/papers/2dargument.html).

5. Dependencies

The main question at issue here is what it takes for an object to have a property. This is intended in the 'proximal' sense. Obviously, there can be long chains of causation that lead to an object's coming to have a certain property. But the issue here is what is the immediate requirement for an object to have a property. I am interested here in properties which have a dependency of some sort on other properties. Presumably there are also 'basic' properties which objects have primitively or directly and presumably the most fundamental properties are among (or perhaps exhaust) the set of basic properties.

Dependency Type 1. An object, a, can have property F in virtue of other, completely distinct, objects possessing other properties and standing in certain relations. For example, something's having the property of being Canadian currency depends upon other individuals having certain properties. Properties exhibiting Type 1 dependency are extrinsic or relational. We can roughly formalize this, leaving various niceties[9] aside, as:

$$\textbf{DT1} \quad Fa \rightarrow (\exists\Gamma)(\exists b_1, b_2, \ldots, b_n)(\Gamma b_1, b_2, \ldots, b_n \wedge \Box_l$$

$$(\forall x_1, x_2, \ldots, x_n, y)(\Gamma x_1, x_2, \ldots, x_n \rightarrow Fy))$$

where Γ stands for the properties/relations which underlie the property in question and $\Box_l$ represents full-on logical necessity. Type 1 dependency as defined is extreme insofar as it ignores any contribution of the object, a, itself. The property of 'being accompanied in the universe' would seem to be an example of pure Type 1 dependency (which immediately generates infinitely many others of essentially the same ilk, *e.g.* being accompanied in the universe with a red thing, *etc.*). Impure forms represent combinations of the types of dependencies listed here.

Dependency Type 2. Fa holds in virtue of a possessing some other property, G. We can call such properties derivative. All determinable properties are derivative from the determinate possessed (*e.g.* the property of being colored holds in virtue of the object having some definite color). Logical 'inclusion' more generally provides other examples of derivative properties (*e.g.* the property of weighing at least 9 kg derives from the property of weighing 10 kg, given that the object in question weighs exactly 10 kg). We might try to

[9] For example, the explicit clause asserting that each b is different from a. In addition, here and below I ignore outer universal quantification for simplicity of presentation.

formalize the idea of F being derivatively dependent on G thus:

DT2 $Fa \rightarrow (\exists G)(G \neq F \wedge Ga \wedge \Box_l(\forall x)(Gx \rightarrow Fx))$

Dependency Type 3. Fa holds in virtue of the constituent parts of a having distinct properties and participating in certain relations. Call such properties 'logically emergent'. It is crucial to note that in this sort of dependency nothing is required for a to possess F other than for its parts to have the properties in question, where these parts are evolving and interacting according to the laws applicable to those parts. For example, it suffices for something to have the property of being transparent if its constituent molecules have certain atomic/electronic properties subject to the actual laws of QED. But, obviously, if the relevant properties are hypothetically imagined to be connected in some other way than they are in the actual world, then a may or may not possess F. In a world where the laws of QED were different, perhaps transparency would not emerge. We can, more or less, formalize this notion as follows:

DT3 $Fa \rightarrow (\exists \Gamma)(\exists z_1, z_2, \ldots, z_n)(Cz_1, z_2, \ldots, z_n a \wedge$

$\Gamma z_1, z_2, \ldots, z_n \wedge \Box_l(\forall x_1, x_2, \ldots, x_n, y)$

$(Cz_1, z_2, \ldots, z_n y \wedge \Gamma z_1, z_2, \ldots, z_n \rightarrow F_y))$

where 'Cxyz' represents 'x and y constitute z'.

I think Type 3 dependency is the core idea which funds MPD based physicalism. According to the physicalist, all it takes to get mentality into the world is for the physical to be arranged thus-and-so and for the physical to operate according to the actual physical laws. Officially, however, MPD physicalism is not committed to any claims about the constituent structure of objects possessing mental properties. Officially, all that is required is that there can be no change in mental properties without a change in a physical property *somewhere* in the world at issue. That is, the physical-to-mental relation could be a Type 1 dependency (a kind of pure externalism of the mental which is highly implausible) or an amalgam of Type 1 and Type 3 dependency (something like this is what naturalist externalists about the mental have in mind). The intrusion of Type 1 dependency leads to bizarre modal structures in which there are possible worlds that contain zombies but differ physically from the actual world only in the position of a neutrino on the other side of the universe. I suggested that logical transparency would eliminate such modal oddities since the position of the neutrino presumably

would not figure in the deduction of mental properties from the physical ground. We could thus absolutely rule out that the 'shifted neutrino' world is a zombie world.

Dependency Type 4. Fa holds in virtue of the constituents of a having exactly the same property. Call such properties compositional. Mass is an example and, if we take specific mass properties, is actually highly interesting. For example, the specific mass property exemplified by a hydrogen atom is 1.007825037 amu. The mass of a proton is 1.00727638 amu and that of the electron is 0.000548579867 amu. The atomic mass is not quite the same as the sum of the masses of the components because the energy which binds the electron to the proton must be taken into account according to the relativistic principle of energy-mass equivalence. This latter is an empirical law which governs the causal interaction of the proton and the electron. The empirical principle involved could have been different. But, crucially, the principle is a physical law which operates over purely physical properties.

There are many other examples of compositional Type 4 dependency. However, all the examples that come readily to mind of true compositional properties are physically fundamental. Massive things get their mass from the mass of their components, charged things get their charge from the charges of their components, *etc*. It is hard to think of genuine cases of non-fundamental compositional properties. Initially appealing counterexamples to the claim that compositional properties are fundamental arise from almost any 'mass term'. Water is made of water, gold of gold, and so on. However, all such substances resolve themselves into more fundamental structures upon which the identity of the substance supervenes. This observation goes some way in explaining why it is that compositional properties seem to be fundamental.

While Type 4 dependency is, of course, a sub-class of Type 3 and needs no separate formalization, the tendency toward fundamentality of compositionally dependent properties makes Type 4 special.

Dependency Type 5. Fa holds in virtue of its parts having properties or entering relations which *proximally cause* a to have F. Call these properties 'causally emergent'. A possible formalization requires only a typographically minor alteration to that of Type 3 dependency, namely:

$$\textbf{DT5}\quad Fa \rightarrow (\exists\Gamma)(\exists z_1, z_2, \ldots, z_n)(Cz_1, z_2, \ldots, z_n a$$

$$\wedge\ \Gamma z_1, z_2, \ldots, z_n \wedge \Box_n(\forall x_1, x_2, \ldots, x_n, y)$$

$$(Cx_1, x_2, \ldots, x_n, y \wedge \Gamma x_1, x_2, \ldots, x_n \rightarrow F_y))$$

where $\Box_n$ represents nomological or causal necessity. It is important to emphasize that this is supposed to represent the proximal ground for a's having property F (there is no shortage of 'long range' Type 5 dependency).

Although this conception seems to be perfectly consistent and it does not seem impossible that certain properties exemplify Type 5 dependency,[10] there is a big question about whether there actually are any examples of Type 5 dependency as opposed to Type 1–4 dependency. In terms of a theological metaphor, cases of Type 5 dependency would require God to create the stuff forming the constituents of an object (along with the laws governing these constituents) but also and in addition create a new fundamental law which caused the possession of property F when the constituents got into certain arrangements.

One intriguing and often suggested possibility is that quantum entanglement provides a genuine example of Type 5 dependency.[11] It is a fundamental feature of quantum mechanics that any two states of a system can be 'added' to form a new state (this is the linear superposition principle). Given certain other core features of quantum mechanics, this leads to the following possibility. Consider a system which creates particles in pairs with some conserved property, such as spin, which can take either a positive or negative value (for simplicity we can ignore the magnitude and just label these + and −). Assuming we start with zero total spin, then the spins

[10] One philosopher who has argued against the coherence of Type 5 dependence is Galen Strawson who, in 'Realistic Monism: Why Physicalism Entails Panpsychism' (*Consciousness and its Place in Nature: Does Physicalism Entail Panpsychism?*, edited by A. Freeman (Exeter: Imprint Academic, 2006), 3–31), takes the impossibility of Type 5 dependence to rule out any form of genuine emergence of consciousness from a purely physical substrate and thus to support panpsychism (which is, in my view, best understood as a form of Type 4 dependence, about which see below).

[11] Paul Teller can perhaps be read as advocating such a view, although it may be that he is only endorsing the weaker claim that quantum systems exhibit a kind of irreducible ontological holism (see his 'Relational Holism and Quantum Mechanics', *British Journal for the Philosophy of Science* **37** (1986), 71–81). I think that discussions of emergence in quantum mechanics tend to miss the distinction between holism and Type 5 dependency. This is often coupled with the assimilation of Type 3 dependency to 'part whole reductionism'. But while Type 3 dependence is compatible with mereological reductionism it is not equivalent to it, as the example of mass illustrates.

of the individual particles (call them A and B) must sum to zero. There are two ways that this can happen, namely, if A has spin + and B has spin −, or the reverse. There is no way to control the polarity of a particular particle's spin during pair production, so when, for example, A is created it is in a superposition of + and − spin, and similarly for B. But since the total spin of the system has to be zero, if we measure A and find it has spin + then we immediately know that B has spin − (or at least will, if measured, give a guaranteed result of −).

This state, known as the singlet state, can be expressed in this highly non-standard but hopefully perspicuous form:

$$\text{Singlet } \frac{1}{2}(A^{+} \otimes B^{-}) + \frac{1}{2}(A^{-} \otimes B^{+})$$

where the ⊗ symbol represents the 'joint state' of particle A and B. There is no way to decompose this complex superposition into a form in which the A states and B states are separated, hence the use of the term 'entanglement' to describe such states.

Entanglement has many peculiarities. Notoriously, since measurement in effect forces Singlet into one of its terms, it entails that no matter how far apart A and B are, upon measuring A to be + (−), it is instantaneously fixed that a measurement of B will give − (+). Also, it seems clear that quantum mechanics is endorsing, or perhaps revealing, some kind of holism about entangled states; they are not reducible to purely local states of the component particles. Furthermore, there is no way to devise any local properties of A and B which can 'carry' the observed correlations between possible spin measurements (though this is possible if the properties can exchange information instantaneously across any distance). So entanglement is very weird.

However, our question is about whether this phenomenon gives any support to the claim that modern science provides examples of Type 5 dependence. The answer seems clearly to be 'no'. The superposition principle and the formation of joint states are fundamental principles of basic quantum physics. Singlet is a state fully described by the fundamental laws of quantum mechanics. Just as the complexities of mass composition depend logically upon the laws governing mass and energy (which are basic physical laws), so too the complexities of entanglement are a logical consequence of the laws governing the basic properties and interactions of the quantum world. In some ways, the analogy is quite close. One would be severely misguided to think that one could simply add up the masses of the constituents of an object, considered independently, to compute the total mass

of the object. The mass of an object is, in a sense, not reducible to the mass properties of the individual components. Their interaction has to be taken into account. But this does not show that an object's mass has a Type 5 dependency on the mass of its constituents, because the interactions are governed by the fundamental laws at the level of the constituents themselves. Similarly, the entangled state is a predictable logical consequence of the basic laws of the quantum particles and their interactions. In fact, it *was* predicted, by Schrödinger, who introduced the term 'entanglement', at the birth of quantum mechanics. There is no hint of Type 5 dependency here, although the oddity of entanglement is also emphasized by this analogy – unlike in the case of mass, there is apparently no (current ongoing) interaction between the entangled particles!

Although modern science does not accept the existence of Type 5 dependence, it nonetheless does not seem to be impossible and has been seriously endorsed. A recent example is that of 19th-20th century emergentism. Once we allow for the possibility of causal emergence, then the step from the constituting microstructure and interactions to the macrophysical property might actually fail to fall under our Type 3 dependency relation. Many of the so-called British emergentists mentioned above believed this, and used chemistry as the lynchpin and, as it seemed to them, utterly uncontroversial example of a sort of emergence which could not be predicted nor was logically determined by sub-chemical physical processes. On their view, chemical, and many other including mental, properties stand precisely in a Type 5 rather than Type 3 dependency relation to the underlying physical structures and interactions. If the causal emergence of chemistry was actually true, there would be possible worlds where the underlying physics was the same as in our world but in which chemistry was different (or absent). Such chemical-zombie worlds are obviously an analogue to the mental zombie worlds under discussion here. They arise from the fact that basic laws of nature are not necessary but rather vary from possible world to possible world (non-basic laws are relatively necessary since they follow logically from basic laws.)

Following these precedents, a dualist of a certain emergentist bent will insist on Type 5 dependency of the mind on the physical, that there are fundamental non-physical laws of emergence which regulate the generation of mental features when certain physical configurations come into being.

Is there really a difference between Type 3 and Type 5 dependence? Type 3 is intended to express the idea that there is a logically necessary connection between F and Γ (using the latter symbol, as

above, as an abbreviation for the complex of properties and relations which underlie F). It is thus an intrinsically stronger claim than the one made by Type 5 dependency. If the connection described in DT3 holds between F and Γ, then there are no possible worlds where something has constituents with properties/relations Γ but lacks F. But in the case of the connection described by Type 5 dependency, all it takes to break the connection is a breakdown in causal laws, and all causal laws are non-necessary. There are possible worlds where something has constituents which stand in the appropriate relation but lacks F if the connection is only of Type 5.

Now, consider once again the dualist who – in the concessionary spirit of generous moderation – allows that physicalism might be true, that there are possible worlds in which physicalism is true. Physicalism asserts that the relation between mental properties and the physical is of Type 3.

In light of this we now face once again the question of how the concessionary dualist should understand the nature of dualism. If it is accepted that physicalism is true at some worlds, then at those worlds there is a Type 3 dependence of the mental on the physical. The dualist might think to embrace this possibility by postulating worlds in which there are non-physical structures which similarly support a Type 3 dependency relation of the mental. This is akin to the frequently advanced idea that a functionalist about the mind is indifferent about the 'stuff' which realizes mental properties. In Hilary Putnam's words, 'strictly speaking, a Turing Machine need not even be a physical system'.[12] As Putnam develops this idea, he skates over the delicate question of what would be the nature of the realizing states of a 'Cartesian mind' which was a Turing Machine. But it is obvious that if the realizing stuff is not intrinsically mental then we do not have a form of dualism in the relevant sense. In a world where a non-physical and non-mental substrate provides the ground of a Type 3 dependency, the mental is not a fundamental part of that world's ontology. This is a dualism which divides the physical from the unknown non-physical *and* non-mental substrate which realizes the mental, not the desired dualism of mind and body.

Now, it is possible to imagine a dualist who accepted both that the realizing substrate and the realized properties were mental, while also allowing that there could be a Type 3 dependence from the physical to the mental. But in such a case the mental nature of the realizing

[12] H. Putnam, 'The Mental Life of Some Machines' in *Mind, Language and Reality: Philosophical Papers, Vol. 2* (Cambridge: Cambridge University Press, 1975), 412.

substrate would be irrelevant to the realized mental properties. What would be relevant are the causal relations amongst the structures of the realizing substrate which can be perfectly mimicked by the structure of physical states in possible worlds where physicalism is true, and these, by hypothesis, do not depend in any way on the substrate being mental. Such a view would be a bizarre form of panpsychism, but one in which the mental nature of the subvening components would be irrelevant to the mental nature of the macro-mental features which supervene upon them. Once again, we would not have a mind-body dualism but a dualism of the physical and 'causal structure' (already familiar from standard functionalism).

More sensibly, a dualist could maintain that the mental nature of the constituents was essential for the generation of the macro-mental features. That is, such a dualist would hold that the dependency relation between the subvening constituents and supervening mental states was of Type 4 (a sub-class of Type 3). Thus it could be held that mental properties are 'composed' out of the mental properties of an object's constituents in a way analogous to the composition of the mass of an object from the masses of its constituents. As noted above, the dependency relation involved with mass is not a simple addition but involves empirical laws with some definite and complex content, albeit laws which operate at the level of the constituents. Similarly, for the dualist, principles operating at the level of the 'micro-mentality' of the object's constituents would generate its 'macro-mental' state.

Our concessionary dualist has been forced to recognize that mental features cannot have a Type 3 dependence on non-mental properties (or non-mental aspects of notionally mental properties). What about the other dependency types? Type 4 is a variant of Type 3, but if we are not allowed mental properties at the base, Type 4 cannot solve the problem. It is evident that mentality is not a logical derivative of being physical in the way required for Type 2 dependency.[13] In any case, as noted above, Type 2 dependency is, from the point of view of the physicalist, merely a disguised form of Type 3 dependency since the macrophysical property required for a Type 2 dependence would itself be dependent on the microphysical via a Type 3

[13] Stephen Yablo has offered what I take to be a surprisingly fruitful *analogy* between the mind-matter relation and the determinable-determinate relation ('Mental Causation', *The Philosophical Review* **101** (1992), 245–80). There are problems with trying to take the idea literally, however (for some, see E. Cox, 'Crimson Brain, Red Mind: Yablo on Mental Causation', *Dialectica* **62:1** (2008), 77–99).

relation. Nor does the mental seem to be an extrinsic or relational property, still less a relational property dependent on the non-mental. Worse, as noted, this threatens to let zombies in by the back door so to speak. If mentality is extrinsic then zombies could exist in a world with a few atoms moved just a few nanometres from their actual positions.[14] Another obvious worry here is that if the physicalist can specify the region required to subvene the mental, then the state of this region ought to underpin a Type 3 determination of the mental within that region. This leaves only Type 5 for the dualist to embrace. But Type 5 dependency of the mental on the physical leads directly to the possibility of zombies.[15]

If this is what dualism involves, then it turns out that the concessionary physicalist is *also* in trouble. If there is a possible world where dualism is true then there is a world, a different world which is specified by the dualist relative to the dualist world allowed by the physicalist, which is an MPD of the actual world but which lacks mentality. This the physicalist cannot allow. So there can be no concessionary physicalists after all.

If one is to be a physicalist, one must regard dualism as absolutely impossible (hence inconceivable). If one is to be a dualist, one must regard physicalism as absolutely impossible (hence inconceivable). For a physicalist, it seems there is no coherent notion of the non-physical, save for the merely verbal concession that some nominally non-physical 'stuff' could mimic the causal relationships enjoyed by the physical which are themselves sufficient to logically determine mentality. The dualist intuition is that there is a coherent idea of a different realm of being: consciousness, which is not logically dependent on the physical (even if it is as a matter of fact causally dependent on the physical). Though this gives every appearance of

[14] Leaving aside the 'shifted particle' worlds, it is an explicitly noted, if unfortunate, consequence of the view that mental properties are extrinsic that a creature that failed to stand in the requisite relations would thereby qualify as a zombie, no matter how physically (hence behaviorally) similar it might be to one of us; see F. Dretske, *Naturalizing the Mind*, (Cambridge, MA: MIT Press, 1995). This seems almost a *reductio* of such accounts of the mental, with the added 'bonus' of strongly supporting the general intuition of the possibility of zombies.

[15] Of course, there remain a host of dualist views in which the mental bears no relation of determination upon the physical (pretty much all the usual suspects: occasionalism, parallelism, *etc*.). But this sort of dualist will have no reason to deny that there are worlds which are MPDs of the actual world that lack mentality. Zombies remain possible.

both conceivability and possibility, the physicalist dare not make such a concession.[16]

University of Toronto

[16] This paper began as a commentary on Alexandru Manafu's presentation at the 2008 CPA meeting and I would like to thank Alexandru for sparking my reflections on this topic. Additional comments of David Chalmers have been very helpful, though of such a generous extent that they could not all be addressed here.

The Representational vs. the Relational View of Visual Experience

BRIAN P. MCLAUGHLIN

Abstract

In *Reference and Consciousness*,[1] John Campbell attempts to a make a case that what he calls 'the Relational View' of visual experience, a view that he champions, is superior to what he calls 'the Representational View'.[2] I argue that his attempt fails. In section 1, I spell out the two views. In section 2, I outline Campbell's case that the Relational View is superior to the Representational View and offer a diagnosis of where Campbell goes wrong. In section 3, I examine the case in detail and argue that it fails. Finally, in section 4, I mention two very well-known problems for the Relational View that are unresolved in the book.

1. The Two Views

By the Representational View of visual experience, Campbell seems to have in mind the following four-fold view: (a) what it is for an experience to be a visual experience is for it to have one of the members of a family of highly determinable phenomenal characters; (b) visual experience is common to visual perceptual experience and visual hallucinatory experience in that a visual perceptual experience and a visual hallucinatory experience could have exactly the same completely determinate phenomenal character; (c) whether a visual experience is perceptual or hallucinatory depends on how it came about; and (d) the phenomenal character of a visual experience, whether the experience is perceptual or hallucinatory, is, or is constituted by, its representational content.[3]

Theses (a)–(c) comprise the traditional view of visual experience, a view that, it is fair to say, remains the dominant view in the United States. Given (b), the view is a 'common factor' view: it entails that a visual perceptual experience and a visual hallucinatory experience can share a common experiential factor. Common factor views contrast with disjunctive views, according to which a visual perceptual

[1] John Campbell, *Reference and Consciousness* (Oxford: Clarendon Press, 2002).

[2] *Ibid*., chs. 6–7.

[3] Hereafter, just to try to avoid prolixity, I'll sometimes drop 'or is constituted by'.

doi:10.1017/S135824611000010X

experience and a visual hallucination share no common experiential factor. As will emerge shortly, the Relational View is a disjunctive view.

Some theorists embrace (a)–(c), but deny (d). Thesis (d) is entailed by so-called representational theories of the phenomenal characters of visual experiences.[4] Proponents of such theories try to justify them, in part, by appeal to the well-known phenomenological thesis of the transparency (or diaphanousness) of visual experience.[5] That thesis has no canonical formulation. I'll use the following formulation: what it is like for one to have a visual experience is for it to be to one as if one is presented with a scene. This what-it-is-like aspect of a visual experience is its phenomenal character. According to the Representationalist, that phenomenal character is or is constituted by the representational content of the experience. Representational theories can differ markedly in their account of what it is for an experience to have a representational content. But they agree that representational contents are semantic contents, in that they are satisfied or instead fail to be satisfied. A visual experience is veridical if its representational content is satisfied; non-veridical if its representational content fails to be satisfied. And its representational content is satisfied just in case matters are as having the visual experience presents matters to the subject as being. A view is a Representational View if and only if it entails (a)–(d). It is not my concern here whether the Representational View is correct. My concern is only with whether Campbell has succeeded in making a case that the Relational View is superior to it.

In the remainder of this section, I'll examine Campbell's presentation of the Relational View in detail, using a scalpel. The aim of the examination is to pinpoint the fundamental, essential difference between the Representational View and the Relational View.

Discussing a case of having a visual perceptual experience as you look around a room, Campbell tells us:

> On a Relational View, the phenomenal character of your experience, as you look around the room, is constituted by the actual layout of the room itself: which particular objects are there,

[4] See, *e.g.*, G. Harman, 'The Intrinsic Quality of Experience' in J. Tomberlin (ed.), *Philosophical Perspectives* (Northridge, Calif.: Ridgeview, 1990); F. Dretske, *Naturalizing the Mind* (Cambridge, Mass.: MIT Press, 1995); W. Lycan, *Consciousness and Experience* (Cambridge, Mass.: MIT Press, 1996), and M. Tye, *Consciousnes, Color, and Content* (Cambridge, Mass.: MIT Press, 2000). Campbell cites the Dretske book and the Tye book (146).

[5] See the references in note 4.

> their intrinsic properties, such as colour and shape, and how they are arranged in relation to one another and to you. On this Relational View, two ordinary observers standing in roughly the same place, looking at the same scene, are bound to have experiences with the same phenomenal character. For the phenomenal character of the experience is constituted by the layout and characteristics of the very same external objects. We have the ordinary notion of a 'view', as when you drag someone up a mountain trail, insisting that he will 'enjoy the view'. In this sense, thousands of people might visit the very same spot and enjoy the very same view. You characterize the experience they are having by saying which view they are enjoying. On the Relational picture, this is the same thing as describing the phenomenal character of their experiences.[6]

A few pages later, he elaborates, saying:

> [T]he Relational View says only that the qualitative character of conscious experience is constituted by the characteristic layout of the objects one is seeing. It is consistent with that to say that only certain of their characteristics constitute one's experience of them. For example, hidden characteristics of the objects will play no role in constituting one's experience of them. Hence, the egocentric spatial layout of the scene may play a role in constituting the qualitative character of one's experience of the scene.[7]

It is Campbell's considered position that the egocentric layout of the scene a visual perceiver sees indeed constitutes the phenomenal (or qualitative) character of the perceiver's visual experience of it.

In the chapter following the one from which the above quotes are taken, Campbell says:

> On the Relational View I discussed in the last chapter, the phenomenal content of the experience of an ordinary observer is constituted by the qualitative character of the view the observer is currently enjoying: which objects and properties are there in the scene, together with the viewpoint from which they are being observed.[8]

[6] J. Campbell, *Reference and Consciousness*, *op. cit.*, 116.
[7] *Ibid.*, 120.
[8] *Ibid.*, 146.

He writes in the passage of 'the phenomenal content of the experience', but he doesn't mean here a kind of content that is satisfied or that fails to be satisfied depending on whether matters are as the content characterizes them as being. He uses 'phenomenal content' here just as a synonym for 'phenomenal character'. Campbell says nothing further, in addition to what he says in the passages above, by way of explicating the notion of a view or scene, or by way of explicating the notion of the egocentric or viewpoint relative layout of the objects in a scene. Although much can be said about these matters, I'll simply assume here that the basic idea of a scene is clear enough for present purposes. I will take it, then, that the Relational View that he takes to be superior to the Representational View is the view that the phenomenal character of a visual perceptual experience is the scene that the perceiver sees, and so consists of the objects that the perceiver sees, some of their properties, and how they are spatially arranged at the time in question relative to a spatial position suitably occupied by the perceiver at that time.

It is common ground between the Representational View and the Relational view that one has a visual perceptual experience if and only if one sees a scene. On the Relational View, the phenomenal character of a visual perceptual experience is the scene that the perceiver sees. The phenomenal character of a visual experience is what it is like for the subject to have the experience. Where the experience is a visual perceptual experience, the Relationalist maintains that what it is like for the subject to have the experience is what the scene experienced is like from the viewpoint occupied by the perceiver, its layout from that viewpoint. The Representationalist will agree, but will say that that is so because the perceptual experience is (completely) veridical. The phenomenal character of the experience is its representational content, and its representational content is satisfied by the scene.

It is common ground between the Representational View and the Relational View that when one has a visual (completely) hallucinatory experience, there is no scene that one sees. But Representationalists maintain that visual hallucinatory experiences have phenomenal characters. When one has a visual hallucinatory experience, it is like something for one to have the experience. What it is like to have a (vivid) visual hallucinatory experience is for it to be as if there is a scene before one. According to the Representationalist, visual hallucinatory experiences have representational contents, and those contents are their phenomenal characters. It is open for a Relationalist to hold that as well. Indeed, on a charitable reading, Campbell, a proponent of the Relational View, holds that. Consider how Campbell

compares and contrasts a case of visually perceiving a dagger with a case of visually hallucinating a dagger. He says:

> Suppose that a dagger is hanging in the air before you, and you are looking at it closely. You are visually attending to it. What can we say to compare and contrast this with the case in which you are having a vivid hallucination of a dagger, and this hallucination is occupying your attention? Just to be fully explicit, the case I have in mind is one in which the ordinary case of attention to a dagger is matched as closely as possible by the hallucinatory experience. That is, if the ordinary dagger seems heavy and substantial, so too does the hallucinatory dagger; the hallucination does not, for example, shimmer unduly, or seem any more blood-stained than daggers usually do.[9]

I think Campbell writes loosely when he writes of 'the hallucinatory dagger'. One can hallucinate a dagger, but there are no such things as hallucinatory daggers. Campbell would, I believe, agree with that claim. The Representationalist claims that what makes an hallucination an hallucination of a dagger is its representational content. Campbell uses the locution 'having a visual experience as of something being F' in characterizing visual hallucinatory experience.[10] When one visually hallucinates a dagger, one has a visual experience as of a dagger. Moreover, Campbell writes of an hallucinator 'seeming to see objects'; thus, he says: 'Even if I am hallucinating, the objects I seem to see…'.[11] Someone visually hallucinating a dagger, 'seems to see' a dagger, which, I take it is intended as just another way of saying that the hallucinator has a visual experience as of a dagger. What Campbell seems to have in mind by saying that a seen dagger 'is matched as closely as possible by the hallucinatory experience [of a dagger]'[12] is that the dagger matches, or satisfies, the general representational content of the hallucination of a dagger. The Relational View, as Campbell presents it (and that is all that I have to go on), is perfectly compatible with the Representationalist account of the nature of hallucinatory experience.

What Relationalists and Representationalists disagree about is the nature of visual perceptual experience. Campbell often states the disagreement this way: on the Relational View, but not on the Representational View, if one has a visual perceptual experience of

[9] *Ibid*., 117.
[10] *Ibid*.
[11] *Ibid*., 121.
[12] *Ibid*., 117.

an object, then that object is a constituent of the experience. Thus, in discussing the difference between visually perceptually experiencing a dagger and visually hallucinating a dagger, he says, for instance:

> On the Relational View…[i]n the case in which there is a dagger, the object itself is a constituent of your experience. The experience is quite different in the case of hallucination, since there is no object to be a constituent of your experience.[13]

And, as concerns the Relational View, he says: '[t]he idea is that visual experiences are relational: the object perceived is a constituent of the conscious experience itself'.[14] Given that the object is a constituent of the experience, it is not (no state of it is) a cause of the experience. Campbell notes, though, that the object will be a cause of a brain state of the subject.[15]

This contrast, however, doesn't capture the essential difference between the Representational View and the Relational View. The reason is that it is open to a Representationalist to maintain that when one visually perceptually experiences an object, the object is a constituent of the experience. Visually perceptually experiencing an object is an experiential relation that a perceiver bears to the object – it is the relation of seeing the object. The object is a relatum of that relation. The Representationalist maintains that one bears that relation to an object by having a visual experience with a certain representational content and the object's bearing an appropriate causal connection to that visual experience. That leaves open, however, whether the visual perceiving of the object is identical with the visual experience with the representational content or is instead distinct from it, but generated by it.[16] The Representationalist can go either way on that issue. Which way a Representationalist goes will depend on where he or she stands on the issue of whether (to use a phrase of Jonathan Bennett[17]) identity holds across the by-relation. The issue is familiar from action theory. Suppose that S kills X by fatally shooting X (by shooting X in a way that appropriately results in X's death). On the view that identity holds across the by-relation, the killing is identical with the shooting.

[13] *Ibid*.
[14] *Ibid*.
[15] *Ibid*, 147.
[16] The *locus classicus* of this notion of generation is A. Goldman, *A Theory of Human Action* (Princeton: Princeton University Press, 1970).
[17] J. Bennett, *Events and Their Names* (Cambridge: Hackett Publishing Company, 1988).

On the alternative view, the killing and the shooting are distinct actions, but the latter generates the former in the circumstances in question (circumstances in which the shooting appropriately brings about the death of X). The issue turns on general problems about the individuation of states and events. A Representationalist can go either way on the issue. Thus, it is open to a Representationalist to maintain that if one visually perceives an object, the object is a constituent of the relational experience of visually perceiving it. The object is, then, not a cause of the perceptual visual experience, though it is, on this view, a cause of a visual experience with a representational content.

As concerns the Relational View, Campbell says:

> It would not be unreasonable to call this view 'naïve realism'. The point of calling it that is to say that on this view, the relation 'S perceives O' is taken as primitive: it is not to be analyzed in some such terms as 'O causes S to have an experiential content as of something's being G'.[18]

It doesn't, however, get at the essential difference between the views to say that the Relationalist holds that 'S perceives O' is primitive, while the Representationalist maintains that it is analyzable in such terms as 'O causes S to have an experiential content as of something's being G'. An analysis requires non-circular necessary and sufficient conditions. It is wide open to a Representationalist to maintain that we cannot state non-circular necessary and sufficient conditions for someone's perceiving an object. As Campbell notes, Representationalists standardly maintain that if S (visually) perceives O, then (i) O exists, (ii) S has an experience with a certain experiential content, and (iii) O (or a state of O) figures as a cause of that experience. But while that is true, (i)-(iii) do not jointly suffice for S's perceiving O. It is a familiar point that there are many objects whose states figure as causes of an experience yet are not objects of the experience. A Representationalist need not maintain that there are non-circular conditions that are individually necessary for, and jointly sufficient for, S's perceiving O. Think of how truly few concepts there are for which there are non-circular necessary and sufficient conditions. Representationalists reject disjunctivism. But they need not embrace conjunctivism, the view that S's perceiving O can be analyzed as the conjunction of certain conditions.[19]

[18] J. Campbell, *Reference and Consciousness*, *op. cit.*, 118.

[19] A similar point is made in M. Johnston, 'The Obscure Object of Hallucination', *Philosophical Studies* **120:1–3** (2004), 113–83.

The fundamental, essential difference between the Representational View and the Relational View is, rather, this: on the Representational View, one visually perceptually experiences an object by having a visual experience with a representational content and the object's bearing an appropriate causal connection to that visual experience. On the Relational View, that is false. Indeed the Relationalist denies that having a visual perceptual experience of an object involves having a visual experience with a representational content. As Campbell puts it:

> On the Relational View, in contrast [to the Representational View], it makes no sense to ask how the subject is representing what she sees.[20]

If the Relational View is right, then a visual perceptual experience could not involve a visual experience that has a phenomenal character in the sense in which an hallucinatory experience has a phenomenal character. We are left with only a disjunctive notion of the phenomenal character of a visual experience. The phenomenal character of a visual perceptual experience is a scene. The phenomenal character of a visual hallucinatory experience is a representational content.

Although Campbell is a Relationalist, he thinks that seeing an object involves psychological factors. These psychological factors include cognitive processing. Alluding to the vision science thesis that the visual system employs feature maps to solve binding problems, he says:

> There is a sense in which cognitive processing is a 'common factor', an element found in both veridical and hallucinatory processing. Whether or not there is an external object being seen, the same features may be located on just the same feature maps, and they may be bound together in just the same way.[21]

He further points out that the cognitive processing in question is information processing, involving cognitive states that have content. And he notes that this invites the view that:

> Since the cognitive contents function as common factors, the experiential contents will be common factors too, exactly the same whether the external object exists or not.[22]

[20] J. Campbell, *Reference and Consciousness*, *op. cit.*, 156.
[21] *Ibid.*, 118.
[22] *Ibid.*

But he holds that although the cognitive contents function as common factors, no experiential contents do. Indeed, he holds that there is no experiential factor that is common to perception and hallucination. To repeat: the Relational View is a disjunctive view.

Let's turn now to why Campbell thinks the Relational View is superior to the Representationalist View.[23]

2. An Outline of Campbell's Case and a Diagnosis

Campbell maintains that in adjudicating the dispute between the Relational and the Representational View,

> the only way to proceed is to ask why we need the notion of the phenomenal character of experience. We have to look at the role the notion plays in our reflective thinking, we have to ask what the point is of the notion.[24]

The common core notion of the phenomenal character of an experience is this: the phenomenal character of an experience is what it is like for the subject to have the experience. Campbell is precisely right about the way to proceed. He takes the role the notion plays in our reflective thinking to be an explanatory role; and that is right too. We invoke the notion for various explanatory purposes; we invoke it in order to explain certain things. If the Relational View can reveal how the phenomenal character of experiences can serve the explanatory purposes in question, and the Representational View cannot do so, then that would be a compelling reason indeed in favor of the Relational View over the Representational View.

As may already be apparent, however, it is curious indeed that Campbell asks why we need the notion of the phenomenal character of experience, given that he is committed to the view that the notion covers two different sorts of things. The phenomenal character of a visual perceptual experience is a scene; the phenomenal character of a visual hallucinatory experience, a kind of representational

[23] I should note that Campbell maintains that some of his objections to Representationalism are objections as well to the disjunctive theories in J. McDowell, 'Singular Thought and the Boundaries of Inner Space' in J. McDowell and P. Pettit (eds), *Subject, Thought, and Context* (Oxford: Oxford University Press, 1986), 137–68; and in W. Child, *Causality, Interpretation, and the Mind* (Oxford: Oxford University Press, 1994).

[24] J. Campbell, *Reference and Consciousness*, *op. cit.*, 120.

content. What reason is there to think that there is an explanatory role that such a disjunctive notion plays?

In fact, Campbell goes on to ask and then to try to answer a question that is only relevant if one takes the phenomenal character of an experience to be a scene, and so to be something that a visual hallucinatory experience lacks. Although he says that we should proceed by asking and then answering the question why we need the notion of the phenomenal character of experience, he instead proceeds to ask and answer a question that the Representationalist regards as a very different question. He proceeds to pursue the question of why we need the notion of 'an experience of an object'. By an experience of an object he means conscious attention to an object, in the sense of one's conscious attention being focused on an object.[25] Thus, he says: 'we use the notion of experience of an object – or more precisely, conscious attention to an object'.[26] More precisely still, he means one's conscious visual attention being focused on a seen object – a visual perceptual experience of an object in which, as he puts it, the object is 'highlighted' in the experience by one's focus of attention.[27] One has a visual perceptual experience of an object if and only if one sees the object. One can see an object without one's conscious visual attention being focused on it. But Campbell has in mind attentively seeing an object. Hereafter, I'll follow Campbell in using 'experience of an object' in this way, as short-hand for one's conscious visual attention being focused on an object that one sees. In the present context, I'll also follow him in sometimes using 'conscious attention to an object' as short-hand for that; and I'll sometimes use 'visual attention to an object' in that way too.

On the Relational View, the phenomenal character of a visual perceptual experience is the scene that the perceiver sees. Moreover, on the Relational View, the phenomenal character of an experience of an object is the object itself and certain of its properties, all of which are elements of the scene that the subject sees. Thus, on the Relational View, the question of why we need the notion of an experience of

[25] This is a state. In contrast, focusing one's attention on an object is a mental action. We're concerned with the state, not the action.

[26] J. Campbell, *Reference and Consciousness*, *op. cit.*, 138. It may well seem redundant to speak of 'conscious attention'. But Campbell speaks of 'conscious' attention, rather than simply of attention, to distinguish 'attention as a phenomenon of consciousness' from 'attention as an information-processing phenomenon' of the sort described in vision science (3). In this paper, I am trying to employ his terminology as often as I can.

[27] *Ibid.*, 10.

an object is one of the questions we must answer to answer the question why we need the notion of the phenomenal character of an experience. From the Representationalist's point of view, however, the question why we need the notion of an experience of an object is a very different question from the question of why we need the notion of the phenomenal character of experience. On the Representational View, one can have an experience with a phenomenal character, even when one fails to experience any object at all (as in a case of complete hallucination).

Campbell cites, by my count, basically four explanatory roles that he maintains the notion of an experience of an object is supposed to play. Here they are, presented mainly in his own words:

> *Explanatory role 1*: 'Experience of objects has to explain how it is that we can have the conception of objects as mind-independent'.[28]
>
> *Explanatory role 2*: '[E]xperience of objects… explains our ability to think demonstratively about perceived objects'; 'Whatever else is true of it, experience of objects has to explain our ability to think about those very objects'.[29]
>
> *Explanatory role 3*: '[E]xperience of objects has a role to play in explaining our knowledge of reference' … '[O]ne's experience [of an object] can explain one's knowledge of the reference of a demonstrative'.[30]
>
> *Explanatory role 4*: 'There are certain basic patterns of inference involving demonstratives whose correctness cannot be grasped by someone interpreting the demonstrative by means of conscious attention to the object, if "conscious attention" is conceived on the common factor model'.[31] Experience of objects must explain how we grasp the correctness of the basic patterns of inference in question.

One issue is whether experience of objects is supposed to explain all or indeed any of the matters in question. A second issue is whether, even if it is, what that is supposed to have to do with the dispute between Relationalism and Representationalism. In section 3, I'll address the first issue at length. For the remainder of this section, I'll focus on the second issue.

[28] *Ibid*., 121.
[29] *Ibid*., 114.
[30] *Ibid*., 115 and 149.
[31] *Ibid*., 129.

Campbell thinks that the Representational View can show us how experience of an object can play explanatory roles 1–4 only if it can show us how an experience's having a certain representational content can play those roles. He thinks that because he thinks this:

> The Representationalist is committed to saying that it is in virtue of its representational content that experience can play its explanatory role.[32]

Campbell offers no justification whatsoever for this claim, however. He just assumes it. In fact, the Representationalist has no such commitment. Indeed the Representationalist will flatly reject the claim that whenever an experience explains something, it does so in virtue of its representational content. The Representationalist will flatly reject that, even if the Representationalist holds the thesis of identity across the by-relation. Even such a Representationalist holds that the fact that an experience is an experience of a certain object explains certain matters, and that the fact that the experience has a certain representational content explains certain other matters. The Representationalist maintains that the former can explain things the latter cannot explain, because the latter does not entail the former. The fact that an experience has a certain representational content will not entail that the experience is an experience of an object. Thus, from the fact that an experience's being an experience of an object explains something, it does not follow that the experience's having a certain representational content explains that something. Moreover, the Representationalist maintains that to invoke the fact that an experience is an experience of an object in the explanation of matters properly explained by the fact that the experience has a certain representational content would be to invoke superfluous information.

In discussing the explanatory role of experiencing an object, Campbell says at one point: 'experience of the object has to be enough to guarantee that the object exists'.[33] But that is no problem whatsoever for Representationalism. If one experiences O (*i.e.* consciously visually attends to O), then O exists. For one can experience (*i.e.* consciously attend to) only what exists. Instances of the schema 'S experiences O' ('S consciously visually attends to O') entail corresponding instances of the schema 'O exists'.

Here, then, is my diagnosis of where Campbell fundamentally goes wrong in making his case that the Relational View is superior to the

[32] *Ibid.*, 147.
[33] *Ibid.*, 128.

Representational View. He makes the mistaken assumption that according to the Representationalist, whatever an experience explains in virtue of being an experience of an object (or of a certain object), the experience explains in virtue of its representational content. That, as I noted, is a claim the Representationalist flatly rejects, on the ground that the fact that an experience has a certain representational content does not entail that the experience is (or generates) an experience of an object. Consider an experience of a certain object, and let us suppose, as Representationalists typically do, that that experience itself has a representational content, rather than being generated by an experience with a representational content. According to the Representationalist, the experience of the object will explain certain things in virtue of being an experience of that object, and the experience will explain certain other things in virtue of having a certain representational content. Being an experience of a certain object and being an experience with a certain representational content play different explanatory roles; they are invoked for different explanatory purposes. That is directly relevant to evaluating Campbell's case against the Representational View. Even if the notion of an experience of an object plays explanatory roles 1–4, and the notion of an experience's having a representational content does not play any of those roles, that fact alone would be no reason to favor the Relational View over the Representational View.

Still, though, suppose that experience of an object is supposed to play one or more of the explanatory roles specified in 1–4. If experiencing an object plays the role on a Relationalist conception of it, but fails to play to it on a Representationalist conception of it, then that would be a reason in favor of Relationalism over Representationalism. So, in the next section, I'll examine 1–4 in detail. And I'll argue that Campbell fails to identify any such reason for favoring Relationalism over Representationalism.

3. Would-Be Explanatory Roles 1–4 Examined

Consider, again, explanatory role 1: 'Experience of objects has to explain how it is that we can have the conception of objects as mind-independent'.[34] It is a tall order indeed for experience of objects to explain that. It seems that many kinds of animals can focus their conscious visual attention on objects they see (think of an animal visually tracking its prey), yet are such that they neither

[34] *Ibid*., 121.

have a conception of objects as mind-independent nor, for that matter, a conception of objects as mind-dependent; indeed such a conception seems cognitively closed to them. That is certainly so if having a conception of objects as mind-independent requires having a conception of mind. But it is uncertain whether it does. Campbell does not elaborate on what he means by the conception of objects as mind-independent.

Campbell tells us:

> On the common factor view, all that experience of the object provides you is a conscious image of the object—the image which bears the representational content… We cannot exact the conception of a mind-independent world from a mind-dependent image.[35]

The Representational View, however, is a kind of common factor view, and the above remarks are a serious mischaracterization of it. On the Representational View, an experience of an object relates one to an object, not to an image of an object. Indeed, it is no part of the Representational View that experiences of objects even involve images. Also, as I noted, it is even open to a Representationalist to maintain that an experience of an object includes the object as a constituent (if the Representationalist rejects the thesis of identity across the by-relation). Further, let's say that one indirectly sees an object if and only if one sees the object by seeing something else; and that one directly sees an object if and only if one sees the object, but not indirectly. The Representationalist, like the Relationalist, maintains that we can directly see objects.[36] To be sure, on the Representational View, one sees an object by having an experience with a certain representational content. But of course one doesn't see the experience; the experience doesn't look any way to one; moreover, the experience isn't an object of visual attention. That is so even in cases of complete hallucination. In such cases, the subject doesn't see anything; nothing looks anyway to the subject; there is nothing that is an object of visual attention. It is just that it is experientially for the hallucinator as if there is an object before him.

The Representationalist, you will recall, embraces the phenomenological thesis of the transparency of visual experience. What it is like to have a visual experience is what it is like to view a scene. As I

[35] *Ibid*.

[36] I use 'object' in a broad sense here, so that the facing physical surface of an object counts as an object.

mentioned earlier, Representationalists typically appeal to the transparency thesis to help justify the view that the phenomenal character of a visual experience is a representational content. Although Campbell never mentions the transparency thesis by name, it is clear from his discussion that he takes it to be at least relevant to explanatory role 1. Again calling a visual experience in the Representationalists sense an 'image', Campbell says:

> You might have thought that the immediate response of a common factor theorist to this argument is that the image provides the conception of an objective world simply by displaying the world as objective. Even if I am hallucinating, the objects I seem to see, seem to be mind-independent objects; and surely that is all that is needed.[37]

I take it that he has in mind here that a common factor theorist such as a Representationalist can appeal to the transparency thesis. He readily acknowledges that the Representationalist can make the above reply. But he says:

> The problem with this reply is that it takes for granted the intentionality of experience. That is, it takes for granted that experience of the world is a way of grasping thoughts about the world. To see an object is, on this conception, to grasp a demonstrative proposition.[38]

It is true that, on the Representational View, experiences have intentionality or aboutness: a visual experience, whether perceptual or hallucinatory, will have a representational content, and a representational content is a kind of intentional content. But it is not part of the Representational View that experience is a way of grasping thoughts, or that to see an object is to grasp a demonstrative proposition. Indeed, the leading Representationalists deny both that experience is a way of grasping thoughts and that to see an object is to grasp a demonstrative proposition.[39] Grasping thoughts and grasping demonstrative propositions are conceptual affairs. While we normally exercise concepts while having a visual experience, the leading Representationalists maintain that the representational content of an experience that is its phenomenal character is a nonconceptual content.[40] To take a stock example from the literature,

[37] J. Campbell, *Reference and Consciousness*, *op. cit.*, 121.
[38] *Ibid.*
[39] See the references in note 4.
[40] See the references in note 4.

something can look red and at a right-angle to one, even when one lacks both the concept of redness and the concept of a right angle. The Representationalist can hold that in virtue of their non-conceptual contents, visual experiences, whether perceptual or hallucinatory, 'display the world as objective' (to use Campbell's phrase).

The fact that Representationalists can (and typically do) maintain that the relevant representational contents are non-conceptual contents bears on what Campbell takes to be the difficulty that Representationalism has in explaining how experience of an object can achieve explanatory role 2. I'll turn now to that matter. Explanatory role 2 is, you will recall, this:

> '[E]xperience of objects … explains our ability to think demonstratively about perceived objects.' 'Whatever else is true of it, experience of objects has to explain our ability to think about those very objects'.[41]

Campbell says:

> if you think of experience as intentional, as merely one among many ways of grasping thoughts, you cannot allow it this explanatory role … if all there is to experience of objects is the grasping of demonstrative thoughts about them, then experience of objects is just one among many ways in which you exercise your conceptual skills.[42]

As I noted, however, although Representationalists maintain that experiences are intentional, the leading Representationalists deny that they are ways of grasping thoughts. To repeat: they maintain that the representational contents of experiences that are their phenomenal characters are non-conceptual contents.

Campbell is aware that a Representationalist can maintain that the representational contents in question are non-conceptual contents. He mentions that possibility in connection with a specific proposal of Tyler Burge.[43] I won't pause here to discuss Burge's proposal. For present purposes, the first point to note is just that Burge posits a demonstrative element in visual experience, and Campbell notes that 'a "demonstrative element" can be as it is, whether or not it refers to an object, and independently of which object it refers

[41] J. Campbell, *Reference and Consciousness*, *op. cit.*, 114.
[42] *Ibid.*, 122.
[43] T. Burge, 'Vision and Intentional Content' in E. Lepore and R. van Gulick (eds), *John Searle and his Critics* (Oxford: Blackwell, 1991).

to'.[44] That point is certainly correct. The second point to note is that he acknowledges that it might be maintained that

> the demonstrative element is not to be regarded as something that is itself immediately involved in thought about the object; it belongs to a category of perceptual representation that is more primitive than thought, and therefore it can play an explanatory role here.[45]

He immediately goes on to say, however:

> But the move to thinking in terms of 'non-conceptual' content does not help. All that is within the perceiver's subjective life is the demonstrative element itself. The aspects of the content which fix the reference of a particular demonstrative element on a particular occasion are not themselves to be assumed to be available to the subject. The thing that is subjectively available – the demonstrative element – cannot of itself, therefore, distinguish between presentation of one object and presentation of another. Nor can it, of itself, provide an assurance that the demonstrative element refers at all.[46]

It is certainly true that the demonstrative element (if such there be) cannot of itself distinguish the presentation of one object from the presentation of another, and that it cannot of itself even guarantee that it has any referent at all. What, if any, referent it has will be determined by the relevant contextual factors. But a point that I have been belaboring is that a Representationalist can appeal, in an explanation, directly to the fact that we experience objects – that is, directly to the fact that our visual attention can be focused on an object that we see.

In spelling out explanatory aim 2, Campbell tells us:

> Whatever else is true of it, experience of objects has to explain our ability to think about those very objects. So a characterization of the phenomenal content of experience of objects has to show how it is that experience, so described, can be what makes it possible for us to think about those objects demonstratively.[47]

Here he makes the mistake that I have repeatedly underscored. He mistakenly assumes that the Representationalist is committed to the view that whatever an experience of an object explains, it explains

[44] J. Campbell, *Reference and Consciousness*, *op. cit.*, 125.
[45] *Ibid.*
[46] *Ibid.*
[47] *Ibid.*, 114.

in virtue of its representational content. To repeat: not only is the Representationalist not committed to that, the Representationalist flatly denies it.

Our ability to experience an object – that is, our ability to consciously visually attend to a seen object – will not by itself explain our ability to think demonstratively about an object we see. The reason is that the latter ability involves the ability to exercise demonstrative concepts in conscious thought, and the former doesn't. The fact that we can consciously attend to seen objects thus won't by itself explain how we can form demonstrative thoughts about seen objects. It will, however, at least figure in the explanation. When one experiences an object – when one's conscious visual attention is focused on a object that one sees – and one uses a demonstrative in conscious thought to try to demonstrate the object that one's attention is focused on, the demonstrative demonstrates that object. I take that to be common ground between the Relationalist and the Representationalist.

Let us turn to explanatory role 3, which is, you will recall:

> '[E]xperience of objects has a role to play in explaining our knowledge of reference'; '[O]ne's experience [of an object] can explain one's knowledge of the reference of a demonstrative'. [48]

I find this would-be explanatory role difficult to assess, because I find it obscure what Campbell means by 'knowledge of the reference of a demonstrative'. One thing he sometimes seems to mean by it is knowing which object is demonstrated by a demonstrative. Knowing which object is demonstrated is a kind of knowing-that. One knows which object is demonstrated when one knows of some object that it is the object that is demonstrated. One's conscious attention being focused on an object that is the referent of a demonstrative could, of course, causally contribute to one's coming to know which object is demonstrated. (So, for that matter, could turning one's head in a certain direction.) But the fact that one's conscious visual attention is focused on an object that is the reference of a demonstrative could not possibly, by itself, explain how one has such *de re* knowledge-that. Still it can figure in the explanation of such a case of *de re* knowledge-that. The main point to note for present purposes is just that whatever explanatory role experience of an object plays as concerns such *de re* knowledge-that, Campbell gives us no reason to think there is any conflict between its playing that role and the Representational View being correct.

[48] *Ibid*., 115 and 149.

Sometimes, however, Campbell seems to mean something different by 'knowlege of the reference of a demonstrative'. He says, for instance, 'Experience of a perceived object is what provides you with knowledge of the reference of a demonstrative referring to it'.[49] Indeed the claim of experience of an object providing knowledge of reference is repeated often in the book. There is no clear sense in which experience of an object that is the referent of a demonstrative 'provides one' with knowledge that it is the reference of the demonstrative. Perhaps, then, Campbell has something different in mind from *de re* knowledge-that. There is a distinction between knowing which thing is demonstrated by a demonstrative and knowing the thing which is demonstrated by a demonstrative. The former is *de re* knowledge-that, but the latter is knowledge by acquaintance. Some of Campbell's remarks suggest he might have in mind knowledge by acquaintance. He says, for instance:

> I will argue that only this [Relational] view, on which experience of an object is a simple relation holding between perceiver and object, can characterize the kind of acquaintance with objects that provides knowledge of reference.[50]

By knowledge of reference, he means here knowledge of the reference of a demonstrative. The suggestion I'm now floating is that he doesn't mean knowing which thing is demonstrated, but rather knowing the thing which is demonstrated. He may be claiming that experience of an object provides knowledge by acquaintance of the object. But there are serious reasons to doubt that. If the demonstrated object is, say, the 4,000th cup produced in the year 2,000 in a certain factory in New Jersey, then to have knowledge by acquaintance of the object that is demonstrated is to have knowledge by acquaintance of the 4,000th cup produced in the year 2,000 in a certain factory in New Jersey. One can have knowledge by acquaintance of the cup that is the 4,000th cup produced in the year 2,000 in a certain factory in New Jersey, without having any idea that the cup is the 4,000th cup produced in the year 2,000 in a certain factory in New Jersey, indeed without ever having heard of New Jersey. Similarly, one can have knowledge by acquaintance of an object that is the referent of a demonstrative without having any idea that it is the referent of a demonstrative; indeed one can have such knowledge by acquaintance without even having the concept of the referent of a demonstrative. Given that, I'm inclined to think that he doesn't mean knowledge

[49] *Ibid*., 114.
[50] *Ibid*., 115.

by acquaintance of the thing that is the referent of a demonstrative. But, as I noted above, it is also the case that there is no clear sense to the claim that experience of an object provides *de re* knowledge-that. I confess that I don't really know what Campbell means by 'knowledge of reference'. So, I won't discuss role 3 any further.[51]

Let us turn now to explanatory role 4:

> 'There are certain basic patterns of inference involving demonstratives whose correctness cannot be grasped by someone interpreting the demonstrative by means of conscious attention to the object, if "conscious attention" is conceived on the common factor model'.[52] Experience of objects must explain how we grasp the correctness of the basic patterns of inference in question.

In his discussion of 4, Campbell notes that we can visually track an object over time. And he calls our attention to the following pattern of inference, where the demonstrative 'that' demonstrates a woman that one's conscious visual attention is focused on:

> That woman is running.
> That woman is jumping.
> That woman is running and jumping.[53]

As concerns this pattern of inference, he tells us:

> Recognizing the validity of the inference requires that your experience should make the sameness of the object transparent to you; but, on the common factor conception, that is precisely what your experience of the object cannot do. On the common factor conception, your experience of the object would have been exactly the same whether there was one woman there throughout, or many, or none. So your experience in itself, on

[51] I discuss these matters further, however, in an unpublished manuscript entitled 'Attention and Object'. I am not the only reader of Campbell to be puzzled by what he means by 'knowledge of the reference of a demonstrative'. See S.D. Kelly, 'Reference and Attention: A Difficult Connection', *Philosophical Studies* **120** (2004), 277–86.

[52] J. Campbell, *Reference and Consciousness*, *op. cit.*, 129.

[53] *Ibid*. This is not the happiest of examples, since it cannot be the case that the woman is running and jumping at the same time, even though she may, say, be moving her legs in a running fashion will in flight from a jump. (Think of long distance jumpers.) But we could recast the example in terms of walking and chewing gum.

> the common factor picture, can provide no guarantee of the sameness of the object throughout.[54]

Again, the common factor view conception he has in mind is the Representationalist conception. Campbell makes here the same sort of mistake that I have repeatedly underscored. He mistakenly assumes that if a Representationalist is to explain how we can recognize the validity of such an inference, the explanation must be in terms of the representational content of visual experiences.

On the Representational View, it is indeed the case that two experiences can have the same representational content even if they are experiences of different women. One might be a visual experience of Judy, the other a visual experience of her identical twin sister, Trudy. Indeed a visual experience of Judy could have the same representational content as an hallucinatory experience. But although that is the case, it is not the case that on the Representational View your experience of the woman would have been exactly the same whether there was one woman there throughout, or many, or none. Suppose, again, that the woman that you are visually perceptually experiencing is Judy. Even on a common factor view, a visual experience of Judy is distinct from a visual experience of Trudy, and both are distinct from any hallucinatory experience. You would not visually experience Judy if the woman before your eyes was Trudy, not Judy. And were you completely hallucinating, there would be no object that you are visually experiencing. The key point to note is that we type experiences in many ways. We type them by their phenomenal characters. But we also type them by their objects. An experience of Judy is different from an experience of Trudy, even when the experiences have exactly the same phenomenal character. The reason is that they have different objects.

Of course, one cannot always tell whether one is experiencing Judy, rather than Trudy, or rather than visually hallucinating, and so not experiencing anything at all. But the Relationalist agrees that one cannot always tell. Moreover, the Relationalist and Representationalist can agree that you don't have to always be able to tell such things in order to recognize the validity of the pattern of demonstrative reasoning in the sort of perceptual circumstance that Campbell describes; for the skeptical scenarios can fail to be relevant alternatives.

Let us turn, then, to the question of whether the Representationalist can explain how, in a visual perceptual circumstance, we could

[54] *Ibid*., 130.

recognize the validity of 'That woman is running; that woman is jumping; that woman is running and jumping'. Representationalists can explain it as follows. We can recognize the validity of the pattern of demonstrative reasoning, because we can see that it continues to be the same woman involved in the activities. That, I take it, is what the Relationalist would say too. Indeed, the Representationalist and the Relationalist can offer the same explanation.

We would be able to continue to see that it is the same woman, in part, because we would be able to continue to visually attend to the woman through the interval in question. And we would be able to do so, because our visual system can track an object over time and change of location, as well as over other changes. As Campbell notes our visual system can so track objects as a result of the sub-personal level informational processes by which it solves various binding problems. The Representationalist can appeal to that same explanation. On the Representationalist View, an object is an object one sees in virtue of there being an appropriate causal process initiating with a state of the object and terminating with a visual experience with a representational content. Vision science studies the internal to the visual perceiver stage of such a causal process.

4. Two Well-Known Issues Facing the Relational View

The Representationalist holds that there is explanatory work for the phenomenal character of an experience to do that it won't do on the Relational View. The work concerns illusion and hallucination.

Campbell tells us, you will recall:

> On the Relational View, in contrast [to the Representational View], it makes no sense to ask how the subject is representing what she sees.[55]

In cases of visual illusion, however, one is seeing an object, but the object looks some way that it isn't. The Representationalist maintains that this involves the visual experience representing the object as being some way that it isn't. There are both optical illusions, such as a straight stick's looking bent at the water-line, and psychological visual illusions, such as the Müller-Lyer arrows looking the same length. Despite the fact that he argues that the Relational View is superior to the Representational View, Campbell offers no account of how the Relational View can explain visual illusions of either sort.

[55] *Ibid*., 156.

Optical illusions such as a straight-stick's looking bent at the water-line, and psychological visual illusions such as the Müller-Lyer arrows looking the same length are common to all normal human visual perceivers. They are what Representationalists sometimes call 'normal misperceptions'.[56] Campbell does not discuss such cases of illusion. But he makes a few brief remarks about illusions resulting from idiosyncrasies of a perceiver. Immediately following the last quoted remark above, Campbell says:

> You can ask from which position the subject is viewing the scene, and *you can ask whether the subject is an ordinary observer or if there are idiosyncratic factors affecting the nature of her experience*.[57]

He doesn't elaborate on the italicized remark. And there are only a few, brief, related remarks in the book. Campbell says at one point:

> You may, for example, be looking at the world with a jaundiced eye, so that everything you see seems to have a yellowish cast. In that case your visual experience would not have exactly the same content as the visual experience of an ordinary observer looking at the same scene.[58]

There could indeed be a case in which you see the scene before your eyes, but everything in the scene seems to have a yellowish cast, even though nothing in the scene actually has a yellowish cast. If so, you would be suffering an illusion, rather than hallucinating. As concerns this case of illusion, Campbell tells us it is 'entirely compatible with the Relational View'.[59] And he states:

> It is just a mistake to suppose that the Relational View is undermined by the fact that the idiosyncrasies of the perceiver may affect phenomenal content.[60]

But the point is that one may be seeing a scene, and yet the scene looks some way that it isn't. To accommodate this, it seems that we have to

[56] M. Matthen, 'Biological Functions and Perceptual Content', *Journal of Philosophy* **85:1** (1988), 5–27.

[57] J. Campbell, *Reference and Consciousness*, *op. cit.*, 156 (italicizes are mine.)

[58] *Ibid.*, 119. It should be mentioned that things do not look yellowish to a person with jaundice; rather a jaundiced person takes on a yellowish cast due to the excess of bile pigments in the person's blood.

[59] *Ibid.*

[60] *Ibid.*

posit that the visual perceptual experience has a representational content. I can find in Campbell no alternative explanation of illusions involving idiosyncrasies of the perceiver.

Let's turn to hallucination. The Representationalist purports to be able to explain why a visual hallucination of a scene could seem to a subject just like a visual perception of a scene. The Representationalist maintains that a visual hallucination and a visual perception could have the same representational content, and so the same phenomenal character. It is standard for disjunctivists to point out that from the fact that a visual hallucination of a scene seems to a subject just like a visual perception of a scene, it does not follow that the subject is aware of something that is in common between them (other than their both seeming the same). Disjunctivists are indeed right that that does not follow. But the Representationalist maintains that the best explanation of why it seems that way to a subject is that the subject is aware of something that is in common, namely a representational content. The Relationalist, of course, rejects that explanation. But it is a standard challenge to the Relational View to offer a better explanation. Perhaps the Relationalist can do so. But no explanation is even ventured in *Consciousness and Reference*.

These well-known problems for the Relational View concerning illusion and hallucination are left unresolved in the book; indeed they go largely undiscussed. Given that Campbell tries to argue that the Relational View is superior to the Representational View, his failure to properly address these problems is a serious omission.[61]

Rutgers University

[61] I wish to thank Christopher Hill and Susanna Siegel for comments.

Consciousness: Don't Give Up on the Brain

KENNETH AIZAWA

Abstract
In the extended mind literature, one sometimes finds the claim that there is no neural correlate of consciousness. Instead, there is a biological or ecological correlate of consciousness. Consciousness, it is claimed, supervenes on an entire organism in action. Alva Noë is one of the leading proponents of such a view. This paper resists Noë's view. First, it challenges the evidence he offers from neuroplasticity. Second, it presses a problem with paralysis. Third, it draws attention to a challenge from the existence of metamers and visual illusions.

Extended mind is a hot topic these days. In this literature, one sometimes finds the claim that consciousness does not supervene on just the brain, but on the entire organism and perhaps even its environment. Rather than searching for a neural correlate of consciousness, the advocates of extended consciousness maintain that scientists should search for a broader biological correlate of consciousness. According to Alva Noë, one of the view's boldest advocates,

> It is misguided to search for neural correlates of consciousness – at least if these are understood, as they sometimes are, to be neural structures or processes that are alone sufficient for consciousness. There are no such neural structures. How could there be?[1]
>
> Consciousness is not something the brain achieves on its own. Consciousness requires the joint operation of brain, body, and world. Indeed, consciousness is an achievement of the whole animal in its environmental context. I deny, in short, that you are your brain.[2]
>
> Seeing is an activity of exploring the world, one that depends on the world and on the full character of our embodiment.[3]

[1] A. Noë, *Out of Our Heads: Why You are not Your Brain, and Other Lessons from the Biology of Consciousness* (New York, NY: Hill and Wang, 2009), 185.
[2] *Ibid*., 10.
[3] *Ibid*., 146.

doi:10.1017/S1358246110000032

> It is now clear, as it has not been before, that consciousness, like a work of improvisational music, is achieved in action, by us, thanks to our situation in and access to a world we know around us.[4]

This paper will push back (yet again) against the excesses of extended mind. Here the resistance will take place at three points. The first point, to be presented in section 1, is to rebut one of Noë's arguments that there are no neural correlates of consciousness. In *Out of Our Heads*, Noë proposes that experiments involving the 'rewiring' of ferret brains, carried out by Mriganka Sur and his colleagues, show that the intrinsic character of the brain is irrelevant to explaining the difference between vision and audition. Rather than focusing narrowly on the brain, we should focus on the whole organism. The rebuttal to this consists of showing how Noë either misinterprets or misconstrues the results of these experiments. Section 2 will challenge what we might call Noë's 'strong enactivism'. This is the view that bodily movement is necessary for perception or perceptual consciousness. Contra Noë's strong enactivism, there is solid empirical evidence against the view that consciousness requires the joint operation of the brain, body, and world or that seeing is an activity of exploring the world. This evidence consists of certain experiments and clinical observations involving neuromuscular blockade that reveal perception without bodily movement. Section 3 will challenge what we might call Noë's 'weak enactivism'. Noë claims that 'For mere stimulation to constitute perceptual experience – that is, for it to have genuine world-presenting content – the perceiver must possess and make use of *sensorimotor knowledge*'.[5] Perhaps we can drop the idea that a perceiver must make use of sensorimotor knowledge in the form of bodily movement in order to perceive. Perhaps instead, in the absence of the joint operation of the brain, body, and world, tacit knowledge of sensorimotor contingencies has the role in perception that Noë attributes to it. The contrast between the strong and the weak views comes to this. The weak view requires only possessing a particular body of knowledge, where the strong view requires possessing a particular body of knowledge and putting it to work through active exploration of one's environment or through the joint operation of the brain, body, and world. The difficulty for weak enactivism is that certain simple experiments falsify some of its apparent predictions.

[4] *Ibid*., 186.
[5] *Ibid*., 10.

1. Noë on the Neuroplasticity of 'Rewired' Ferrets

In *Out of Our Heads*, Noë argues that Mriganka Sur's work on ferrets shows that conscious experience does not depend on the intrinsic character of the brain.[6] He thinks this work shows that the brain alone is not the supervenience base for conscious experiences. In these experiments, Sur performed surgical interventions on newborn ferrets. These interventions induce the formation of neural connections between the eye and the medial geniculate nucleus (MGN) which in turn connects to primary auditory cortex, A1. Noë proposes that this procedure transforms a ferret so that area A1 in adults gives rise to visual experiences, rather than auditory experiences, *while leaving A1 itself unchanged*. (See Figure 1). It is because A1 itself is supposed to be unchanged and yet to give rise to visual experiences that we must look beyond A1, and indeed beyond the brain, to explain the visual character of experience. In other words, his goal is to show that the brain by itself is not the supervenience base of conscious states.

As it may be surprising to think that Noë interprets Sur's work in this way, it is worthwhile reviewing Noë's discussion in detail. In doing this, it is important to distinguish the vast majority of what Noë gets right about Sur's work from the important bit that he gets wrong. Early in his discussion of Sur's work, Noë writes:

> Until now, [scientists] have not been able to bridge what is sometimes called 'the explanatory gap' between neural states on the one hand and conscious experience on the other. From my perspective, this is not surprising. The reason we cannot explain the quality of experience in terms of the intrinsic nature of the brain's action is that, in fact, there is nothing distinctively visual about the brain's action.[7]

We can take it as common ground that there is currently an explanatory gap between neural states and consciousness. Noë, however, wants to go further than this and say that the reason there is an explanatory gap is that we are looking in the wrong place for the biological correlate of consciousness. We should not be looking at the brain; we

[6] This conclusion differs from those he and Susan Hurley drew in S. Hurley and A. Noë, 'Neural Plasticity and Consciousness', *Biology and Philosophy* **18** (2003), 131–68. For the purpose of this discussion, however, we may concede these other conclusions.

[7] A. Noë, *Out of Our Heads: Why You are not Your Brain, and Other Lessons from the Biology of Consciousness*, *op. cit.*, 53.

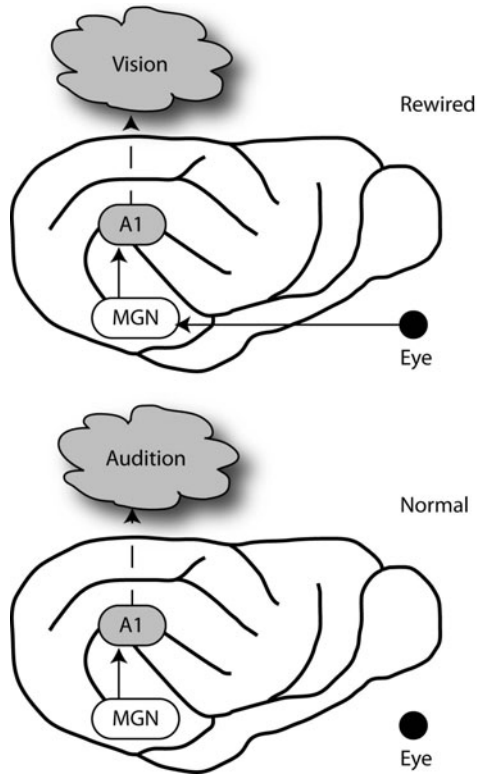

Figure 1. Noë's Interpretation of Sur's Experiments.

should be looking at brain-body-world interactions. We should not be looking for a neural correlate of consciousness; we should be looking for an ecological correlate of consciousness. This argument, however, is not the same as Noë's neuroplasticity argument, since nothing in it depends on neuroplasticity.[8] A bit later, Noë writes:

[8] It is not, however, all that weighty an argument. It appears to be quite common in the course of the history of science to know that some object or structure gives rise to some phenomenon long before having much idea how the object actually gives rise to the phenomenon. For example, it was known for millennia that the sun produces light, but only in the last hundred years or so was it discovered that the sun produces light through nuclear fusion. It was known for thousands of years that muscles were responsible for many bodily movements, but only in recent decades have scientists begun to unravel the biochemical level mechanisms that give muscles their contractile properties. Given that this is standard fare in the history of science, we

> [Sur's work] shows that the link between brain areas and conscious experience (*e.g.* the link between the auditory cortex and auditory experiences and that between the visual cortex and visual experiences) is malleable.[9]

Here is a conclusion that one can accept consistent with believing in a neural correlate of consciousness. One who believes in neural correlates of consciousness can comfortably say that, in Sur's experiments, auditory cortex gives rise to visual experiences. That just means that the tissue that normally gives rise to audition now gives rise to visual experience. This, by itself, does not show that the intrinsic properties of the brain are an inadequate realization base for visual or auditory experiences. For what might be going on is that the surgical interventions convert tissue that once had a structure suitable for realizing auditory experiences into tissue suitable for realizing visual experiences. In fact, as we shall see, this is a fairly plausible interpretation of what actually happens. Next Noë writes:

> The fact that it is possible in this way to vary consciousness in relation to its neural underpinnings teaches that there isn't anything special about the cells in the so-called visual cortex that makes them visual. Cells in the auditory cortex can be visual just as well. There is no necessary connection between the character of experience and the behavior of certain cells.[10]

Here again, we have a conclusion that the advocate of neural correlates of consciousness can accept. One might say that cells from one area of the brain that normally do one task can be reconfigured to perform another task. And what might explain how cells that normally realize auditory experiences come to realize visual experiences has to do with the way in which the cells are interconnected, rather than their individual physiological properties. So, what Noë says to this point is correct, but nothing that challenges the hypothesis that there are neural correlates of consciousness. Noë, however, comes much closer to misleading his readers in the next passage:

---

should not be all that surprised to find solid evidence that the brain gives rise to conscious experiences, even though we do not have much of a clue as to how the brain actually does it.

[9] A. Noë, *Out of Our Heads: Why You are not Your Brain, and Other Lessons from the Biology of Consciousness*, *op. cit.*, 54.

[10] *Ibid.*

> And this finding in turn means that if we want to understand why certain cells or certain brain areas are participating in seeing and not hearing, or in hearing and not seeing, we need to look beyond the immediate neural activity itself. The character of conscious experience can vary even though the neural activity underpinning it does not change. This is the basic lesson of Sur's studies. It follows, then, that what determines and controls the character of conscious experience is not the associated neural activity.[11]

Perhaps one can say that the neural activity is the same, maybe in the sense that both normal and rewired auditory cortex fire normal action potentials. But, it is easy enough to move from this to the idea that the tissue in area A1 is entirely unchanged by the surgical intervention. This is where Noë risks misleading his readers, since if it were true that the tissue in area A1 is entirely unchanged by Sur's intervention, then indeed, one could not appeal to changes in A1 to explain changes in perceptual experiences. If the brain did not change, but the experiences did, then the biological change that is responsible for the perceptual change must lie elsewhere, as in the ferret's connection to the external world. And, in fact, Noë needs the strong conclusion that A1 is entirely unchanged by Sur's intervention in order to have a compelling argument against a neural correlate of consciousness. After all, if A1 were only the same *in some respects*, that still leaves open the possibility that A1 is changed *in other respects* and that these other respects are the ones that are responsible for A1 subsequently giving rise to visual experiences, rather than auditory experiences. Noë is even more misleading when he claims that

> What explains the consequences for consciousness of this sort of rewiring is not the intrinsic character of the neurophysiological changes themselves; it is, rather, the larger setting or context in which these neurophysiological changes occur.[12]

Here it is possible to read 'the neurophysiological changes' as an explicit recognition on Noë's part that there are changes in rewired A1. One might take this to be the clearly charitable reading. If, however, one does this, then Noë's argument against a neural correlate of consciousness collapses for the reason just noted. If Noë accepts that there are neuroanatomical or neurophysiological changes in A1, then that opens the door to those changes being

[11] *Ibid*.
[12] *Ibid*., 56.

what is responsible for converting the tissue in A1 from a supervenience base for auditory experiences into a supervenience base for visual experiences. In fact, if we review Sur's original publications, it will be clear that there are changes in rewired A1 and that they are plausibly construed as changes that would indeed convert A1 into a supervenience base for visual experiences.

The fundamental flaw in Noë's use of Sur's work is that, while the tissue in the ferret's area A1 does not become exactly like normal ferret V1 after the surgical intervention, it does assume many properties that are like normal V1. (See Figure 2). Contrary to Noë's suggestions, if not explicit assertion, rewired ferret A1 becomes something of a hybrid between normal A1 and normal V1 and perhaps has enough features of normal V1 to make it a viable candidate for a realization base for visual experiences. First, consider some similarities between rewired A1 and normal V1. In the first place, it is not merely that rewired A1 responds to light, rather than sound. In addition, the cells in rewired A1 have response properties like those

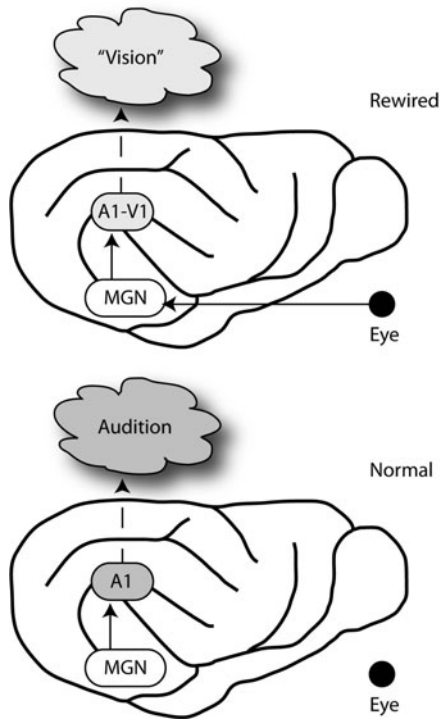

Figure 2. Sur's Interpretation of Sur's Experiments.

in normal V1: some cells are sensitive to oriented lines and bars and others are sensitive to the direction of moving lines and bars. Moreover, the cells in rewired A1 are organized into regions of the sort found in normal V1. As reported in Sur's work, normal A1 has a tonotopic structure in which a given sound frequency corresponds to a single band of tissue running horizontally in A1, where rewired A1 has retinotopic structure in which a given point in the visual field is mapped to a small region of rewired A1.[13] (See Figure 3). These are not obscure details. The abstract to one of these papers includes the following:

> In these [rewired animals], the primary auditory cortex contains a systematic representation of the retina (and of visual space) rather than a representation of the cochlea (and of sound frequency). A representation of a two-dimensional sensory epithelium, the retina, in cortex that normally represents a one-dimensional epithelium, the cochlea, suggests that the same cortical area can support different types of maps.[14]

In addition, other abstracts note that the precision of the tuning of oriented line detectors in rewired A1 is comparable to that of normal V1.[15]

Despite these similarities, there are also differences between rewired A1 and normal V1. First, in rewired A1 the mapping of azimuths is more precise than the mapping of elevations, whereas in normal V1 azimuth is mapped as precisely as is elevation.[16] Second, in rewired A1 there is overlapping projection from MGN to rewired A1 which is unlike the projection from LGN to V1 in normal ferrets.[17] Third, rewired A1 has a less orderly map than does normal V1.[18] Finally, horizontal connections between orientation modules in rewired A1 are less patchy and periodic than are horizontal connections in normal V1.[19]

[13] A. Roe, S. Pallas, J. Hahm and M. Sur, 'A Map of Visual Space Induced in Primary Auditory Cortex', *Science* **250** (1990), 818.

[14] *Ibid*.

[15] See the abstract to J. Sharma, A. Angelucci and M. Sur, 'Induction of Visual Orientation Modules in Auditory Cortex', *Nature* **404** (2000), 841–7.

[16] A. Roe, S. Pallas, J. Hahm and M. Sur, 'A Map of Visual Space Induced in Primary Auditory Cortex', *op. cit*.

[17] *Ibid*.

[18] J. Sharma, A. Angelucci and M. Sur, 'Induction of Visual Orientation Modules in Auditory Cortex', *op. cit*.

[19] *Ibid*.

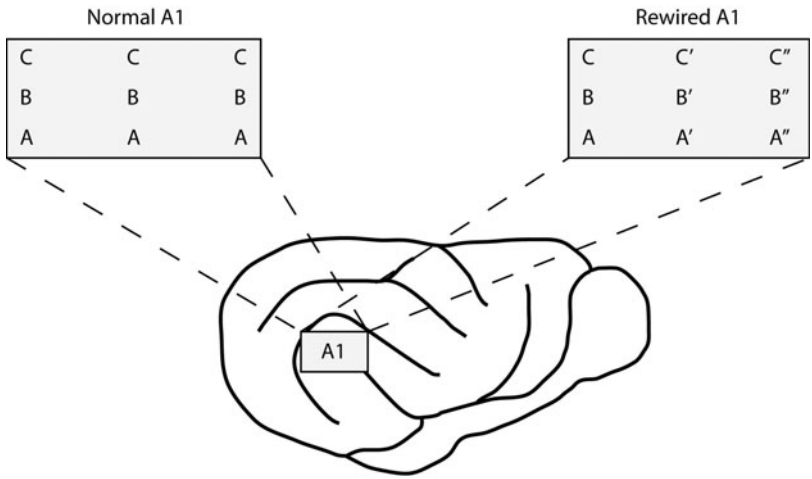

Figure 3. Normal Auditory Cortex (Tonotopically Organized) Versus Rewired Auditory Cortex (Retinopically Organized).

The principal point to take away from this research is that Sur and his colleagues performed certain surgical interventions on newborn ferrets, then raised these ferrets to adulthood in the anticipation that these neonatal interventions will lead to differences in adult area A1. These anticipations have been borne out with Sur and his colleagues documenting the various ways in which rewired ferret A1 comes to resemble normal ferret V1. What these experiments suggest, contrary to Noë's view, is that when we look at the intrinsic features of the brain for a mechanistic explanation of the visual character of experience, we are probably looking in the right place.

2. Strong Enactivism and the Problem of Paralysis

To this point, the discussion has rebutted Noë's analysis of the implications of the rewired ferret experiments. This merely undercuts one of Noë's arguments for his view. Now, however, we shall take a step further and challenge Noë's positive view that we should expect to discover an ecological correlate of consciousness. This will proceed in two stages. The first stage will consider strong enactivism and the second will consider weak enactivism. So, this section will challenge the idea that action in the form of physical movement is necessary for consciousness. In other words, it will provide evidence that conflicts with Noë's view that 'Consciousness requires the joint

operation of brain, body, and world'[20] and that '[s]eeing is an activity of exploring the world, one that depends on the world and on the full character of our embodiment.'[21]

The central problem for Strong Enactivism is the existence of perception despite complete immobilization.[22] There is a broad class of neuromuscular blockers that inhibit neuromuscular function by

[20] A. Noë, *Out of Our Heads: Why You are not Your Brain, and Other Lessons from the Biology of Consciousness*, *op. cit.*, 10.

[21] *Ibid.*, 146. Anthony Morse has noted that action can take the form of mental simulation or imagination, hence that it does not necessarily require physical movement, muscle contraction, or behavior. Moreover, in an unpublished paper, Nivedita Gangopadhyay draws attention to the following:

> Being engaged in a skill does not require that we are at this moment actually executing any action. Just as an explorer is still an explorer when he rests overnight in his tent, just as we are engaged in tying our shoelaces when we briefly pause at one moment in the action, we can be engaged in a sensorimotor skill without at this moment actually performing any action. Nevertheless, we have to be in a state of having cognitive access to the fact that certain movements, if we were to perform them would result in certain consequences in the sensory input. O'Regan, J. K. 'The Sensorimotor Approach to Phenomenal Consciousness' (2006), Retrieved June 18, 2009, from http://www.whatfeelingislike.net/tiki-index.php?page=The+Sensorimotor+Approach+to+Phenomenal+Consciousness.
>
> Neither O'Regan nor I hold that action or behavior is necessary for qualia to occur. What we have argued is that the quality of experience depends on implicit knowledge of the sensory effects of movements (whether or not one actually moves). A. Noë, 'Magic Realism and the Limits of Intelligibility: What Makes us Conscious', *Philosophical Perspectives* **21** (2007), 457–474.

So, there are grounds for thinking that Noë's view is that neither perception nor consciousness require bodily movements. Nevertheless, there are also grounds for thinking that Noë's view is that perception and consciousness *do* require bodily movements, namely, the passages cited. Cf. also, A. Noë, *Action in Perception* (Cambridge, MA: MIT Press, 2004), 1–3.

[22] Locked-in syndrome provides another type of problematic case for Noë. In *Out of Our Heads*, Noë notes that there are individuals who suffer from locked-in syndrome (A. Noë, *Out of Our Heads: Why You are not Your Brain, and Other Lessons from the Biology of Consciousness*, *op. cit.*, 14–16). These individuals appear to be totally unconscious, displaying no actions or behaviors commonly taken to be indicative of consciousness, but they are nonetheless conscious. How is this possible on Noë's view? If consciousness requires the joint operation of the brain, body and

binding to the muscle receptors for the neuromuscular transmitter acetyl choline.[23] Acetyl choline cannot cause normal muscular contraction, since the muscle relaxer blocks its binding sites on muscles. For decades, both experimental and clinical research has shown that neuromuscular blockade induces paralysis, but not the loss of consciousness or perception or awareness. Indeed, muscle relaxers have generally been found not to affect cognitive activity at all. Present purposes will be served by reviewing just one experimental study and a handful of clinical studies.[24]

Prior to administering their subjects the neuromuscular blocker vecuronium, Topulos and his colleagues applied a tourniquet to one arm.[25] This slowed the transmission of the vecuronium from the bloodstream to the neuromuscular junctions, so that the bound arm remained functional. This allowed the immobilized subjects to communicate with the investigators using a pre-established system of finger gestures. During this period, the experimenters asked subjects questions requiring yes or no answers. After the experiment, subjects could also recall events that took place during the paralysis. These results are extremely hard to reconcile with the strong enactivist idea that 'Consciousness requires the joint operation of brain, body, and world'. Consider the details more closely. All subjects perceived or were conscious of the questions asked of them. All subjects

---

world – if seeing is an activity of exploring the world – then how can there be completely inactive individuals who are nonetheless conscious?

[23] This discussion borrows from F. Adams and K. Aizawa, *The Bounds of Cognition* (Malden, MA: Blackwell Publishers, 2008), Chapter 9.

[24] Other studies not reviewed here include S. M. Smith, H. O. Brown, J. E. Toman, and L. S. Goodman, 'The Lack of Cerebral Effects of d-Tubocurarine', *Anesthesiology* **8** (1947), 1–14; E. Campbell, S. Godfrey, T. Clark, S. Freedman and J. Norman, 'The Effect of Muscular Paralysis Induced by Tubocurarine on the Duration and Sensation of Breath-holding during Hypercapnia', *Clinical Science* **36** (1969), 323; E. Campbell, S. Godfrey, T. Clark, J. G. Robson and J. Norman, 'The Effect of Muscular Paralysis Induced by Tubocurarine on the Duration and Sensation of Breath-holding', *Clinical Science* **32** (1967), 425–32; A. Froese and A. Bryan, 'Effects of Anesthesia and Paralysis on Diaphragmatic Mechanics in Man', *Anesthesiology* **41** (1974), 242–55; J. Stevens, R. Emerson, G. Gerstein, T. Kallos, G. Neufeld and C. Nichols, 'Paralysis of the Awake Human: Visual Perceptions', *Vision Research* **16** (1976), 93–8.

[25] G. Topulos, R. Lansing and R. Banzett, 'The Experience of Complete Neuromuscular Blockade in Awake Humans', *Survey of Anesthesiology* **38** (1993), 133.

found the process of tracheal intubation, which was part of the purpose of the study, to be 'extremely uncomfortable'. They were conscious of discomfort. In addition, all complained of the bitter taste of the lidocaine spray that was used to facilitate the intubation. Two of the subjects worried about the discomfort of the procedure. Two subjects explicitly stated that they did not know when paralysis had set in until they tried to move. Presumably, on Noë's theory, consciousness should fade exactly when paralysis sets in, since loss of movement is loss of part of the minimal supervenience base for consciousness. But, that did not happen. As a summary, the investigators report that

> Complete neuromuscular block caused no observable impairment of consciousness, sensation, memory, or the ability to think and make decisions. Objective evidence supported this assertion, as subjects responded promptly to questions. When the experimenter misunderstood their answers, subjects recognized this and made a correction. Subjects successfully used a questionnaire with many branching points to communicate their needs. Subjects also accurately recalled specific events that occurred in the room while they were paralyzed. This unimpaired mental function is consistent with the reports of previous investigators.[26]

It is, of course, true that the failure to paralyze one arm does not lead to complete paralysis. Yet, it is unclear why the ability to move one arm (even were this ability exercised) would be the kind of sensorimotor skill that would be sufficient to maintain the apparent capacity to perceive bitter tastes, the discomfort of intubation, the questions of the investigators, and the anxiety over the procedure.

This experimental work is complemented by a large body of clinical work. Anesthesiologists have used muscle relaxers to enable surgeons to better control the body position of patients. In addition, muscle relaxers are useful in suppressing involuntary motor responses, such as gagging, that occur during intubation. An unfortunate side effect of the use of muscle relaxers, however, is that it makes it much more difficult to determine whether a patient is adequately anesthetized. Immobilized patients are able to provide very few signs of the distress they are suffering due to inadequate anesthesia. Consider some of the reports on awareness during surgery. In an anonymous 1979 editorial in the *British Journal of Anaesthesia*, a woman reported in graphic detail her recollections of being aware

[26] *Ibid*.

during a Caesarian section under general anesthesia and neuromuscular blockade.[27] Although immobilized by the anesthesiologist, she remembers feeling the pain of the incision, hearing her baby crying, and feeling the insertion of a nasogastric tube. In a more recent case study, a 74-year-old woman recalled that during her operation '1) she felt pain during the incision of the abdomen, 2) she heard the operator say, "It is difficult to remove all tumors because the adhesion is very strong," and 3) she remembered someone had been walking around her'.[28] These post-operative recollections were independently confirmed by the surgical staff. Because of the obvious importance of the phenomenon of awareness during surgery, there have been numerous studies of its frequency in hospitals in the United States and Europe. In one study involving 45 patients, all recalled hearing sounds or voices, 33 understood and remembered conversations, 21 had visual perceptions, 29 felt being touched, six recalled moderate pain, and eight recalled severe pain.[29]

Cases of awareness during surgery have their methodological drawbacks as a means of challenging Noë's theory of perception and consciousness. To begin with, these studies rely on patient recollection. Sometimes these putative recollections are rather vague, such as that the patient recognized the surgeon's voice and heard the sound of instruments. Such vague recollections could be recollections of dreaming or pure confabulations. In one prospective study of awareness during surgery, participants were interviewed on multiple occasions.[30] The multiple interviews, in conjunction with the procedures used to obtain informed consent, may have increased the chances of false memories. Despite these limitations, however, there do appear to be cases in which completely immobilized, completely paralyzed patients perceive auditory and visual events during their operations. Moreover, unlike the method used by Topulos and his colleagues, the methodology in these studies does

[27] Anonymous, 'On Being Aware', *British Journal of Anaesthesia* **51** (1979), 711–2.

[28] S. Miura, S. Kashimoto, T. Yamaguchi and T. Matsukawa, 'A Case of Awareness with Sevoflurane and Epidural Anesthesia in Ovarian Tumorectomy', *Journal of Clinical Anesthesia* **13** (2001), 227–9.

[29] D. Schwender, H. Kunze-Kronawitter, P. Dietrich, S. Klasing, H. Forst and C. Madler, 'Conscious Awareness during General Anaesthesia: Patients' Perceptions, Emotions, Cognition and Reactions', *British Journal of Anaesthesia* **80** (1998), 133–9.

[30] R. Sandin, G. Enlund, P. Samuelsson and C. Lennmarken, 'Awareness during Anaesthesia: A Prospective Case Study', *The Lancet* **355** (2000), 707–11.

not leave any of part of the body in a nearly fully functional state. The experimental methods and the clinical prospective studies thus depend on the truth of different sets of auxiliary hypotheses. Therefore, they provide, in some measure, independent methods for determining the effects of paralysis on perception. In short, paralysis does not eliminate consciousness.

3. Weak Enactivism and the Problems of Illusions and Metamers

Noë is clearly aware of the challenge posed by cases of paralysis, and so he offers a possible solution, namely, he backs away from the idea that perceptual experience, hence consciousness, requires exploratory behavior. Instead, experience requires only the possession of tacit sensorimotor knowledge (SMK). In *Action and Perception*, he writes:

> Paralysis is certainly not a form of blindness. But isn't that precisely what the enactive view requires, that the paralyzed be experientially blind? No. The enactive view requires that perceivers possess a range of pertinent sensorimotor skills. It seems clear that quadriplegics have the pertinent skill. … More important, paralysis does not undermine the paralyzed person's practical understanding of the ways movement and sensory stimulation depend on each other. Even the paralyzed, whose range of movement is restricted, understand, implicitly and practically, the significance of movement for stimulation. They understand, no less than those who are not disabled, that movement of the eyes to the left produces rightward movement across the visual field, and so forth.[31]

This seems to be a pretty clear endorsement of what we have called 'weak enactivism'. Rather than perceptual experience requiring the *exercise* of sensorimotor skills through bodily movements or behaviors per strong enactivism, it requires only the *possession* of sensorimotor skills. This is a weaker theory in the sense that one could not exercise a skill one did not possess, but one could possess a skill one does not exercise. One could possess at time t the sensorimotor ability to ride a bike, but not exercise that ability at time t either physically or through mental simulation or imagination.

[31] A. Noë, *Action in Perception*, *op. cit.*, 12. Note Noë's minor mis-step of equating paralysis with quadriplegia. The case of muscular blockade indicates the difference.

It is worth noting that this weakening of the theory is not an *ad hoc* move for Noë. At a number of points he invokes this weaker theory to explain other important perceptual phenomena, most notably, amodal completion.[32] To see this, we need some further elaboration of Noë's theory. A very simple illustration of amodal completion is useful here: a circular tile which, from a particular perspective, is occluded by a rectangular tile. (See Figure 4). The example is useful, since it is clear and simple, it is (relatively) phenomenologically salient, and its existence is supported by 'objective' psychophysical evidence.[33] Noë would say that, what one *strictly speaking sees* (from position p) is a black rectangle adjacent to a Pac-Man-shaped figure. By contrast, what one *perceptually experiences* (from position p) is a black rectangle occluding a circle.[34] Noë also describes what one strictly speaking sees and what one perceptually experiences in terms of two kinds of perceptual content.[35]

> Perceptual content has a dual aspect. There's the way experience presents the world as being, as it were apart from your perspective. This is one aspect of its content. And there is *the way* the world is presented in experience, a way that always incorporates some reference to how things look or sound or feel from your

[32] In fact, it is Anthony Morse's contention that Noë only ever intends what is here described as the weaker view. That seems to me incorrect, but it is perhaps fair to say that the most extensively argued version of Noë's theory is the weaker possession version.

[33] See, for example, A. Sekuler and S. Palmer, 'Perception of Partly Occluded Objects: A Microgenetic Analysis', *Journal of Experimental Psychology: General* **121** (1992), 95–111.

[34] Noë uses different terms for this distinction at different times. At one point, the distinction is described as being between what one strictly speaking sees and what one perceptually experiences and, at another, as being between what one strictly speaking perceives and what is perceptually present:

> A cat sits motionless on the far side of a picket fence. You have a sense of the presence of the cat even though, strictly speaking, you only see those parts of the cat that show through the fence. How is it that we can in this way enjoy a perceptual experience as of a whole cat?
>
> These are instances of the problem of perceptual presence. We have a sense of the presence of that which, strictly speaking, we do not perceive. (A. Noë, *Action in Perception*, *op. cit.*, 60).

For simplicity and clarity, the current discussion will use only the former terminology.

[35] A. Noë, *Action in Perception*, *op. cit.*, 163–4.

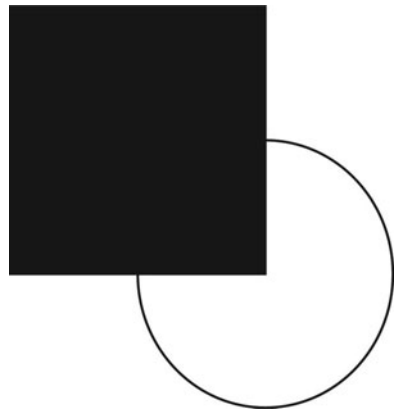

Figure 4. An Instance of Amodal Completion of a Circle Occluded by a Square.

> vantage point. So, for example, your experience presents you with the circularity of the plate, but also with the elliptical shape it presents from here.[36]

Thus, in our amodal completion example, experience 'aperspectivally' presents the world as having a black rectangular tile occluding a white circular tile, but experience 'perspectivally' presents the world having a black rectangular tile adjacent to a Pac-Man shaped tile.

Sensorimotor knowledge enters the picture by linking what one strictly speaking sees and what one perceptually experiences. It links perspectival to aperspectival perceptual content. 'How [things] (merely) appear to be plus sensorimotor knowledge gives you things as they are'.[37] This SMK takes two forms: *movement-dependent* and *object-dependent*.[38] Movement-dependent SMK consists of such things as one's knowing how one's own movements away from p lead to changes in what one strictly speaking sees. Thus, as one moves one's head or walks around the two objects, progressively more or less of the circle is revealed or covered. Object-dependent SMK consists of such things as knowing how an object's movements will lead to changes in what one strictly speaking sees of it. Thus, as the occluding black rectangle moves from one place to another, more or less of the occluded circle will be presented.

[36] *Ibid*., 163.
[37] *Ibid*., 164.
[38] *Ibid*., 64–5.

Noë intends to extend this type of analysis of the simple case of amodal completion to an extremely wide range of cases. For example, he claims that one can perceptually experience a whole cat that is partially occluded by the slats of a picket fence or a whole tomato – including its back side – of which we strictly speaking see only the front face. He maintains that we perceptually experience the color and fine detail of the physical world in the periphery of our visual field, although we do not strictly speaking see it.[39] What links what we strictly speaking see to what we perceptually experience is SMK:

> The ground of this accessibility is our possession of sensorimotor skills. … In particular, the basis of perceptual presence is to be found in those skills whose possession is constitutive, in the ways I have been proposing, of sensory perception. My relation to the cat behind the fence is mediated by such facts as that, when I blink, I lose sight of it altogether, but when I move a few inches to the right, a part of its side that was previously hidden comes into view. My sense of the perceptual presence, now, of that which is now hidden behind a slat in the fence consists in my expectation that by moving my body I can produce the right sort of 'new cat' stimulation.
>
> In this way, we can explain our sense of the perceptual presence of, say, the whole tomato. Our perceptual sense of the tomato's wholeness – of its volume and backside, and so forth – consists in our implicit understanding (our expectation) that movements of our body to the left or right, say, will bring further bits of the tomato in view. Our relation to the unseen bits of the tomato is mediated by patterns of sensorimotor contingency. Similar points can be made across the board for occlusion phenomena.
>
> In general, our sense of the perceptual presence of the detailed world does not consist in our representation of all the detail in consciousness now. Rather, it consists in our access now to all of the detail, and to our knowledge that we have this access. This knowledge takes the form of our comfortable mastery of

[39] In truth, Noë seems a bit equivocal on this score. On the one hand, there is the third paragraph of the cited text below in which Noë claims that we have a perceptual experience, or sense the perceptual presence, of the detail of the world (in the periphery?), but, on the other, he also denies that we experience detail in the periphery: 'we don't experience the periphery of our visual field in anything like the clarity, detail, or focus with which we can take in what we are directly looking at' (*Ibid*., 49).

> the rules of sensorimotor dependence that mediate our relation to the cat and the bottle.[40]

Noë's theory is clearly meant to have a very broad scope.

The general form of the problem for Noë's weak enactivism is that SMK does not appear to mediate between what one strictly speaking sees and what one perceptually experiences. SMK does not appear to form a bridge between 'aperspectival' perceptual content and 'perspectival' perceptual content. So, to begin to unravel Noë's theory, let us return to the black rectangle that, from perspective or position p, occludes a white circle. As is widely known, this perceptual experience occurs even when there is *not* a circle behind the black rectangle. If the part of the white figure that one strictly speaking sees is consistent with a circle, but the occluded portion is highly irregular, (from position p) one still experiences the white portion as circular. How can this be on Noë's theory? In this case, movement will *not* reveal a circle. Movement of one's head, the black rectangle, or the white irregular object will only reveal an irregular object. One cannot have SMK that such and such movements will reveal a circle, since no movements will reveal a circle. In such cases, SMK cannot connect what one strictly speaking sees, a black rectangle adjacent to a Pac-Man shaped figure, to the perceptual experience of a black rectangle occluding a circle, since there does not exist the appropriate SMK.[41]

The 'mistakes' in amodal completion, among other examples, suggest that we should not take Noë to be speaking of 'knowledge' in the epistemologist's sense of something like justified true belief (or what have you), but of 'knowledge' in the psychologist's sense

[40] *Ibid*., 63.

[41] One can construct a similar counterexample by using a bit of text in which a central fixation point is in sharp focus, but in which the periphery is defocused. The entire text appears to be sharp, but this cannot be because one has access now to detail in the periphery (cf. the third paragraph from Noë cited above), since there is no such detail in the periphery.

Notice as well that the problem with amodal completion has been developed in terms of three dimensional objects, such as thin circular and square blocks, rather than in terms of two dimensional drawings that are typically used to illustrate the phenomenon. The problem, however, seems even more serious when we note that there is amodal completion with drawn figures. The problem is apparently harder, since movement-dependent and object-dependent motion SMK does not provide the kind of information that would inform amodal completion. This point is made in B. Nenay, 'Four theories of amodal perception', *Proceedings of the 29th Annual Conference of the Cognitive Science Society* (2007), 1333.

of something like belief or expectation. Moreover, this type of belief is not just any kind of belief or expectation. Noë apparently means the kind that is formed through long practice, and trial and error. It might be acquired only through extensive exploration of, and interaction with, the environment. One of Noë's discussions on this score is the way in which SMK is modified in Ivo Kohler's experiments with inverting and distorting goggles.[42] Very briefly, in these experiments one dons one or another form of distorting lens goggles which immediately proves to be disorienting. With extensive training over the course of several weeks, however, one comes to learn to compensate for the alterations, thereby restoring near normal levels of behavioral competence. It is this kind of sensorimotor knowledge or tacit understanding that Noë probably has in mind.

This refinement or clarification, however, does little to address the fundamental problem with Noë's view. Even repeated exposure to the fact that an occluded object is not a circle appears not to undermine the perceptual experience of a circle. The refractory character of amodal completion is, in fact, a familiar feature of many visual illusions. Consider the Müller-Lyer illusion, the Poggendorff illusion, or the Herman grid illusion. (See Figure 5). Repeatedly measuring the lines in the Müller-Lyer illusion does not make the lines look to be the same length. Repeatedly masking and unmasking the arrowheads and tails in the Müller-Lyer illusion does not make the lines look to have the same length once they are unmasked. Repeatedly aligning a straightedge with the diagonal line of the Poggendorff illusion does not make the apparent break go away when the straightedge is removed. Scanning around the Herman grid over and over does not make the smudges at the intersections go away once they are again viewed in the periphery. The simple sorts of learning that Noë describes with the inverting goggles do not suffice to make these illusions go away.

Take another category of problem cases: metamers. Consider the rectangular grid, G, and uniform gray patch, U, shown in Figure 6. At position p – a distance of about 20 feet or more – G and U look alike. If, however, one moves closer to a position p* – a distance of about 2 feet – G and U look different. G looks like a grid, where U looks like a uniform gray rectangle. Imagine protracted interaction with, and exploration of, G and U. Even with all this experience with the effects of movement on the perspectival appearance of G

[42] See A. Noë, *Action in Perception*, *op. cit.*, 7–10. Noë also discusses Paul Bach-y-Rita's experiments with Tactile-Visual Sensory Substitution (TVSS) equipment in this connection (26–7).

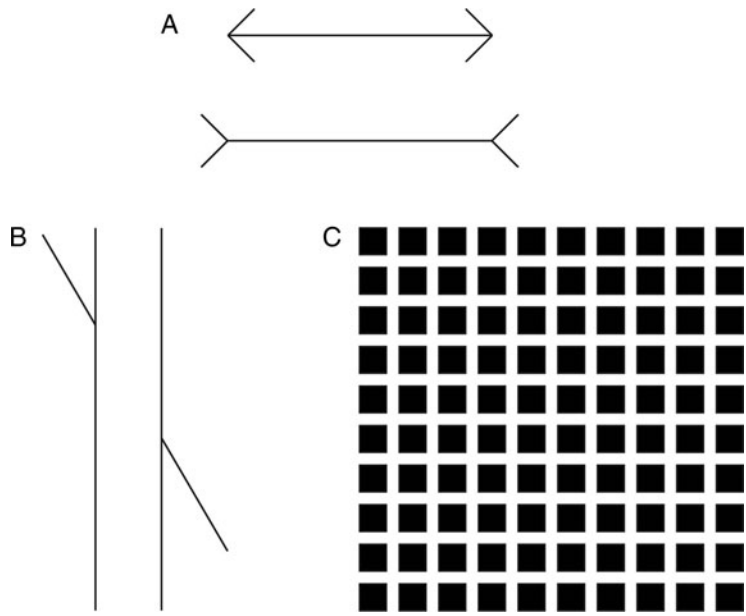

Figure 5. A) the Müller-Lyer Illusion, B) the Poggendorff Illusion, and C) the Hermann Grid Illusion.

and U, when one returns to position p, G and U again look alike. Their perspectival appearance is not connected by SMK to a different aperspectival appearance.[43] In the case of amodal completion, it is supposed to be knowledge of the effects of movement that are supposed to lead a perceiver from how things merely appear perspectivally from position p – as a black square adjacent to a Pac-Man shaped figure – to the way things are aperspectivally – as a black square occluding a circle. The theory does not work in the case of amodal perception, but it clearly does not work in the case of the

[43] In personal correspondence, Anthony Morse has proposed that it is because one knows that U and G look alike from 20 feet away that makes U and G generate the same perceptual experience from 20 feet away, rather than different perceptual experiences from 20 feet away. Making this move to block the counterexample would, however, undermine Noë's entire approach. Presumably in viewing the black rectangle adjacent to a white Pac-Man from position p, one knows that they look like a black rectangle adjacent to white Pac-Man from position p, so that, by parity of reasoning, from position p one should perceptually experience them as a black rectangle and a white Pac-Man, rather than as a black rectangle occluding a white circle.

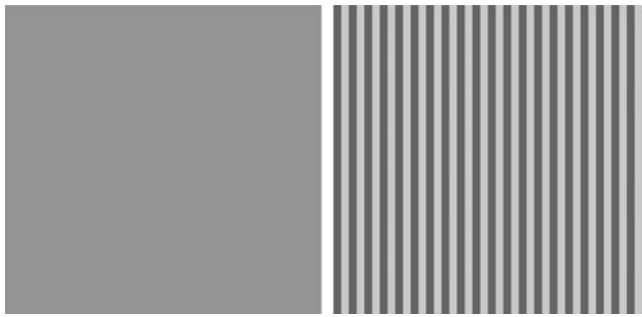

Figure 6. A Uniform Gray Rectangle and a Rectangular Grid.

metameric G and U. SMK again does not take a perceiver beyond the perspectival to the aperspectival.

The problem with G and U appears to generalize quite broadly. Metamers are physical objects or properties (such as U and G) that look the same under a given set of conditions C (such as distance). The recipe for making trouble for Noë is to take two objects M_1 and M_2 that are metameric under conditions C, and then repeatedly vary those conditions so that M_1 and M_2 are no longer metameric. The problem is that when one returns to conditions C, M_1 and M_2 will often remain metameric. Many metamers cannot be unlearned by the kinds of simple sensorimotor learning that Noë describes.

Return now to Noë's claim that perception has a dual aspect. On the one hand, there is the way experience presents the world as being, as when it presents a dinner plate as circular, and on the other hand, there is the way the world is presented from a specific perspective, as when it presents a dinner plate as elliptical. Noë believes that tacit sensorimotor knowledge links the way things look from a particular perspective to how things are:

> perceptual experience is a way of encountering how things are, but it is a way of encountering how things are by making contact with how they appear to be. How they (merely) appear to be plus sensorimotor knowledge gives you things as they are.[44]

The point that illusions and metamers make is that how things appear plus sensorimotor knowledge does *not* give you things as they are. One's tacit sensorimotor knowledge does not always mediate between how things look perspectivally and how things look aperspectivally.

[44] A. Noë, *Action in Perception*, *op. cit.*, 164.

To conclude this section, Noë's theory of the role of tacit sensorimotor knowledge in perception seems to make incorrect scientific predictions about what well-informed viewers will perceive. In many cases of visual illusions, sensorimotor knowledge plus visual stimulation under conditions C 'tells you' not to perceive things that you do perceive. In many cases of metamers, sensorimotor knowledge plus visual stimulation under conditions C 'tells you' to perceive things that you do not perceive. At the risk of overgeneralization, it appears that Noë's theory falsely predicts that there should be no visual illusions or metamers for the well-informed. By Noë's lights, genuine sensorimotor knowledge should let the well-educated perceptually experience the world as it really is, but there is no education *that* good.

4. Conclusion

The core message of this paper is that, for all Noë has argued, it still remains too early to abandon hope for a neural correlate of consciousness. First of all, upon closer examination of the results of Sur's experiments, we can see that the neonatal interventions external to A1 lead to adult differences internal to A1. The general character of the changes is that rewired ferret A1 becomes more like normal ferret V1. This gives us hope that the visual character of visual experience might be explained by properties intrinsic to the brain. Second, Noë's theory that perception requires action of the whole organism founders on the well-documented cases of consciousness in the face of neuromuscular blockade. Third, even the weak theory that perception requires merely the possession, rather than the exercise, of sensorimotor knowledge appears to conflict with certain relatively simple and well-known features of perception, namely, the persistence of visual illusions and metamers in the face of one's sensorimotor understanding of them. In short, for all Noë has said, it is still too early to adopt Noë's holistic approach to consciousness and give up on the brain.[45]

Centenary College of Louisiana

[45] Thanks to Gennady Erlikhman, Nivedita Gangopadhyay, Anthony Morse, and Larry Shapiro for helpful comments on an earlier draft of this paper.

Index of Names

Adams, F. 273
Alexander, S. 110
Alighieri (Dante) 23
Angelucci, A. 270
Aquila, R.E. 24, 33
Aristotle 36, 37, 77, 216
Armstrong, H. 22, 23, 33
Ayer, A.J. 5, 6, 12, 57, 58, 212

Bach-y-Rita, P. 281
Banzett, R. 273
Basile, P. 6, 14, 117, 118
Bayne, T. 130, 167
Bentham, J. 6
Benussi, V. 131
Bergson, H. 109
Berkeley, G. 11, 14, 15, 17
Block, N. 8
Bradley, F.H. 2, 5, 6, 8, 9, 13, 18, 115, 116, 120, 126
Brentano, F. 147
Brown, H.O. 273
Bryan, A. 273
Bühler, K. 131
Byrne, A. 130, 131
Byrne, D. 84

Cameron, von D. 156
Campbell, E. 273
Campbell, J. 239–262
Carnap, R. 170, 211
Carruthers, P. 25, 32
Chalmers, D. 12, 13, 16, 94, 96, 97, 104, 148, 158, 167, 173, 212, 218, 219, 226
Chrysippus 36
Chun, M. 63
Clark, A. 16
Clark, T. 273
Clayton, P. 101
Coleman, S. 74
Collier, J. 68, 73, 95
Cox, E. 235
Crane, T. 149, 158, 177
Cussins, A. 49, 50

Dainton, B. 62, 110, 156, 160, 161, 166, 167, 168, 178, 179, 185, 186, 190, 191, 192
Dennett, D. 10, 34, 47, 48, 64, 203, 204–210
Descartes, R. 2, 9, 16, 45, 53, 54, 55, 56, 57, 60, 64, 65, 71, 73, 81, 83, 146, 227
Detienne, M. 23
De Weerd, P. 63
Dewey, J. 97
Dietrich, P. 275
Diogenes Laertius 36
Douglas, M. 39
Dretske, F. 236, 240

Eccles, J. 225
Eddington, A. 67
Ehrenfels, von C. 131
Einstein, A. 8, 18
Emerson, R. 273
Empedocles 27

Fechner, G. 107
Feynman, R. 78
Fichte, J.G. 71
Flischut, A. 179
Foster, J. 57, 166, 187
Frege, G. 64, 178

Gallagher, S. 176, 178
Garrett, B. 226
Gibson, J.J. 152
Goff, P. 105–6, 112
Goldman, A. 244
Gregory, R.L. 152
Griffin, D.R. 96, 99
Guest, S. 122
Gurwitsch, A. 127

Hadot, P. 34
Hannay, A. 24
Hari, R. 122
Harman, G. 240
Harris, M. 38
Hartshorne, C. 96
Hayward, V. 123
Hegel, G.W.F. 2, 5
Heisenberg, W. 18
Heraclitus 27
Herz, M. 61, 71
Hornung, E. 29, 43, 44
Housman, A.E. 44
Hume, D. 43, 56, 57, 62, 71, 72, 107, 108, 220
Hurley, S. 169, 265
Husserl, E. 2, 9, 141–154, 155, 157, 161, 164, 175, 176, 185, 200

Inwagen, van, P. 75, 76, 77, 78

James, W. 2, 5, 6, 9, 10, 34, 35, 62, 83, 96, 99, 100, 105–110, 115, 116, 120, 126, 161, 164, 175, 176, 183, 185
Joël, K. 145
Johnston, M. 245
Jousmaki, V. 122

Kadowaki, J. K. 32
Kamitani, Y. 122
Kant, I. 2, 14, 45, 52, 53, 54, 55, 56, 61, 62, 71, 81–91, 161
Kelly, S.D. 161, 258
Kim, J. 1, 2, 45–47, 48, 49, 52, 56–60
Klibansky, R. 36
Koffka, K. 131
Köhler, W. 131
Konkle, T. 123

Ladymann, J. 67, 73, 95
Leibniz, G.W. 2, 71, 75, 96, 111, 112, 146
Le Poidevin, R. 157
Levin, D.T. 63
Lewy, H. 38
Lippitt, J. 41
Lloyd, A. 27, 28
Lloyd, D. 157, 174
Locke, J. 64, 101, 161, 185
Lockwood, M. 135
Louth, A. 36

Mabbott, J.D. 190
Mackenna, S. 23
Mai, N. 156
Manafu, A. 237
Marois, R. 63
Martin, M.G.F. 176
Matthen, M. 261
MacDonald, J. 122
McDowell, J. 183, 247
McGurk, H. 121, 122
Meinong, A. 131, 161
McGinn, C. 11, 15, 93, 103, 104
Miller, I. 164, 165
Moore, C. 123
Moore, G.E. 5, 12
Morgan, C. 223
Muir, E. 34

Nāgārjuna 73, 76
Nagel, T. 7, 8, 9, 10, 11, 96, 102, 104, 110, 111, 211
Nelson, J.E. 38, 44
Nozick, R. 71

O'Callaghan, C. 123

Panofsky, E. 36
Papineau, D. 52
Parmenides 73
Pautz, A. 130, 131
Peacocke, C. 151
Peirce, C.S. 18, 96
Perl, E.D. 42
Pessoa, L. 63
Phillips, I. 157, 160, 163, 165, 174, 175, 179
Philo (of Alexandria) 36, 37, 38
Pinker, S. 25
Plato 23, 26, 29, 30, 31, 36, 39, 42
Plotinus 2, 21–44
Pöppel, E. 161
Popper, K. 225
Porphyry 27
Price, R. 130, 131
Prichard, H.A. 156–157
Pythagoras 27

Quine, W.v.O. 6, 7, 18, 79

Redpath, T.H. 6
Reid, T. 155–156, 157, 158, 160, 164, 175, 180
Rist, J. 39
Ryle, G. 9
Robinson, H. 177
Royce, J. 6, 15, 17, 107
Ross, D. 68, 73, 95
Russell, B. 5, 12,13, 25, 67, 68, 95, 100, 162, 212, 213, 226, 227

Santayana, G. 5, 6
Sartre, J.-P. 149
Saxl, F. 36
Schaffer, J. 73, 74
Seager, W. 16, 96, 142, 208, 209, 210
Searle, J. 94–96, 99, 104, 105, 106, 112
Sells, M.A. 42
Shams, L. 122
Shoemaker, S. 64
Siegel, S. 130, 158, 159
Simons, D.J. 63
Simons, P. 102
Shimojo, S. 122
Skrbina, D. 16, 93, 96, 97, 106

Smith, A. 22
Smith, B. 131
Sorabji, R. 27, 28
Spence, C. 122
Spinoza, B. 5, 6, 9, 73, 77, 92, 96, 98
Sprigge, T.L.S. 2, 5–20, 24, 25, 72, 96, 104, 113, 115–118, 119, 120, 121, 124, 125, 126, 130, 132, 133, 138, 139
Spurrett, D. 68, 73, 95
Stout, G.F. 174
Strawson, G. 16, 50, 51, 55, 67, 71, 74, 80, 81, 82, 89, 96, 100, 102, 104, 231
Strawson, P. 45, 53

Thaumasius 27
Teller, P. 231
Tye, M. 169–173, 180, 192, 240

Verhoeven, C. 35
Violentyev, A. 122

Wang, Q. 123
Ward, J. 100
Warren, E.W. 23, 31
Washburn, A.L. 25
Weil, S. 41
Weinberg, S. 74
Wertheimer, M. 131
Whitehead, A.N. 5, 6, 15, 17, 18, 73, 96, 99
Wilson, R. 74
Wisdom, J. 8
Witasek, S. 131
Wittgenstein, L. 9, 97, 214, 215
Wollheim, R. 151, 152

Yablo, S. 235

Zahavi, D. 161, 185
Zampini, M. 122
Zihl, J. 156

The Metaphysics of Consciousness

What is consciousness? What is its place in the physical universe? How did it arise in the course of cosmic evolution? Can there be a genuine scientific understanding of consciousness? And is there such a thing as a genuine human self? These and related questions occupy a prominent place in contemporary studies in metaphysics and philosophy of mind, often involving complex interdisciplinary connections between philosophy, psychology, artificial intelligence, biology, and cognitive neuroscience. At the same time, these questions play a fundamental role in the philosophies of great thinkers of the past such as, among others, Plotinus, Descartes, Leibniz, Kant, William James, and Edmund Husserl. This collection of essays aims at sparking fruitful lines of future inquiry by bringing together systematically oriented historians of philosophy and contemporary leading philosophers of mind to re-examine a broad range of inherited views. In a pluralistic vein—and in the attempt to do full justice to the richness of our mental life—the volume features in-depth examinations not solely of mainstream physicalist doctrines, but also of largely neglected yet in many ways still attractive positions such as Cartesian dualism, idealism, and panpsychism.

Contributors

Kenneth Aizawa, Pierfrancesco Basile, Stephen R. L. Clark, David Cockburn, Barry Dainton, Julian Kiverstein, Geoffrey Madell, Eduard Marbach, Leemon McHenry, Brian P. McLaughlin, Howard Robinson, William Seager, Galen Strawson

CAMBRIDGE UNIVERSITY PRESS
978-0-521-17391-9