THE ROUTLEDGE HANDBOOK OF PANPSYCHISM

"This book provides a rich and novel discussion of one of the most exciting (and hardest) issues in philosophy, namely the nature of consciousness, by taking seriously panpsychism. It contains a large variety and number of quality contributions, both from a historical and a contemporary perspective, which makes it a book of reference indispensable for anyone interested in the field. A great read and an inspiring contribution to the philosophical debate about the nature of consciousness."

Jiri Benovsky, University of Fribourg, Switzerland

"The rise to prominence of panpsychism is a response to a growing disillusionment with orthodox physicalism. If you want to know the history, strengths and weaknesses of this surprising revival of an ancient metaphysics, this wide-ranging collection is an excellent place to start."

Howard Robinson, Central European University, Hungary

Panpsychism is the view that consciousness – the most puzzling and strangest phenomenon in the entire universe – is a fundamental and ubiquitous feature of the world, though in a form very remote from human consciousness. At a very basic level, the world is awake. Panpsychism seems implausible to most, and yet it has experienced a remarkable renaissance of interest over the last quarter century. The reason is the stubbornly intractable problem of consciousness. Despite immense progress in understanding the brain and its relation to states of consciousness, we still really have no idea how consciousness emerges from physical processes which are *presumed* to be entirely non-conscious.

The Routledge Handbook of Panpsychism provides a high-level comprehensive examination and assessment of the subject – its history and contemporary development. It offers 28 chapters, appearing in print here for the first time, from the world's leading researchers on panpsychism. The chapters are divided into four sections that integrate panpsychism's relevance with important issues in philosophy of mind, philosophy of science, metaphysics, and even ethics.

1. Historical Reflections
2. Forms of Panpsychism

First published 2020
by Routledge
52 Vanderbilt Avenue, New York, NY 10017

and by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

Library of Congress Cataloging-in-Publication Data
A catalog record for this book has been requested

ISBN: 978-1-138-81713-5 (hbk)
ISBN: 978-1-315-71770-8 (ebk)

Typeset in Bembo
by Apex CoVantage, LLC

CONTENTS

PART IV
How Does Panpsychism Work?

CONTRIBUTORS

Miri Albahari teaches philosophy at the University of Western Australia. She works in comparative philosophy, on the epistemology and metaphysics of mystical states.

Torin Alter is Professor of Philosophy at The University of Alabama. He writes mostly on consciousness and the mind-body problem

Andrew Bailey is Associate Professor in Philosophy, and Associate Dean (Research & Graduate Studies) for the College of Arts at the University of Guelph, Ontario, Canada. He works on the philosophy of consciousness, embodied cognition and the ethics of AI, as well as the thought of William James.

Pierfrancesco Basile is Lecturer in Philosophy at the University of Bern and at the KSA Luzern, Switzerland. His last book is *Whitehead's Metaphysics of Power: Reconstructing Modern Philosophy* (Edinburgh UP, 2017).

Michael Blamauer worked for several years in the field of philosophy of mind with special focus on panpsychism, notably editing *The Mental as Fundamental: New Perspectives on Panpsychism* (Ontos, 2011).

David Bourget is Associate Professor of Philosophy and Director of the Centre for Digital Philosophy at the University of Western Ontario. He has published a number of articles on consciousness, its role in the mind, and its place in nature.

Monima Chadha teaches philosophy at Monash University in Australia. She works in cross-cultural philosophy with special interest in Abhidharma Buddhism.

David Chalmers is University Professor of Philosophy and Neural Science and Co-Director of the Center for Mind, Brain, and Consciousness at New York University.

Sam Coleman is Reader of Philosophy at the University of Hertfordshire. His main interest is in reconciling scientific accounts of matter and the brain with what we glean from the personal perspective.

Daniel Dombrowski is in the Philosophy Department of Seattle University. Among his many publications is the book *Rethinking the Ontological Argument: A Neoclassical Theistic Response* (Cambridge UP, 2006).

Philip Goff is a philosopher and consciousness researcher at Durham University and author of *Galileo's Error: Foundations for a New Science of Consciousness* (Pantheon, 2019). His research focuses on how to integrate consciousness into our overall theory of reality.

Terry Horgan is Professor of Philosophy at the University of Arizona. He has published (often collaboratively) in philosophy of mind, philosophy of cognitive science, metaphysics, epistemology, and metaethics.

Steven Horst is Professor of Philosophy at Wesleyan University. He is the author of four books, including *Beyond Reduction* (Oxford UP, 2007), *Laws, Mind, and Free Will* (MIT, 2011), and *Cognitive Pluralism* (MIT, 2016).

Robert J. Howell is Dedman Family Distinguished Professor and Chair of the Philosophy Department at Southern Methodist University. His work mostly concerns the nature of subjectivity, the self, and the mind.

Graeme Hunter is Emeritus Professor of Philosophy at University of Ottawa and Research Professor of Philosophy at Dominican University College. He works for the most part on philosophers of the Early Modern Period.

Martin Lin is Professor of Philosophy at Rutgers University in New Brunswick, New Jersey. His research concerns metaphysics and philosophy of mind in the early modern period.

Freya Mathews is Adjunct Professor of Environmental Philosophy at Latrobe University, Australia. She has taught and published very widely in the field of environmental philosophy, particularly in ecological metaphysics.

Angela Mendelovici is Associate Professor of Philosophy at the University of Western Ontario. She has written numerous works on consciousness, intentionality, and the relationship between the two, including her recent monograph, *The Phenomenal Basis of Intentionality* (Oxford UP, 2018).

Hedda Hassel Mørch is associate professor of philosophy at Inland Norway University of Applied Sciences. She was previously a postdoc at Center for Mind, Brain and Consciousness at New York University and University of Oslo.

Yujin Nagasawa is H.G. Wood Professor of the Philosophy of Religion at the University of Birmingham, UK. He is the author and editor of many books including *God and Phenomenal Consciousness* (Cambridge UP, 2008) and *Consciousness in the Physical World* (with Torin Alter; Oxford UP, 2015).

Paavo Pylkkänen is Vice Dean of Research of the Faculty of Arts and Head of Department of Philosophy, History and Art Studies at the University of Helsinki, where he has worked as Senior Lecturer in Theoretical Philosophy. He is also Associate Professor in Theoretical Philosophy at the University of Skövde, Sweden. His research areas are philosophy of mind, philosophy of science, and the foundations of quantum theory. He is the author of *Mind, Matter and the Implicate Order* (Springer, 2007).

Luke Roelofs is a postdoctoral researcher at New York University, working on social cognition and the metaphysics of consciousness. His new book, ***Combining Minds*** (Oxford UP, 2019), address the combination problem for panpsychism.

William Seager is Professor of Philosophy at the University of Toronto Scarborough. His most recent book is *Theories of Consciousness* (2nd ed; Routledge, 2016).

Michael Silberstein is Professor of Philosophy at Elizabethtown College and Affiliated Faculty in the philosophy department at the University of Maryland, College Park, where he is also a faculty member in the Foundations of Physics Program and a Fellow on the Committee for Philosophy and the Sciences. His most recent book is *Beyond the Dynamical Universe: Unifying Block Universe Physics and Time as Experienced* (Oxford UP 2018).

David Skrbina is a senior lecturer in philosophy at the University of Michigan at Dearborn. His research interests include panpsychism, philosophy of technology, and environmental philosophy. His books include *Panpsychism in the West* (revised edition, MIT Press, 2017), *The Metaphysics of Technology* (Routledge, 2015), and most recently, *The Jesus Hoax* (Creative Fire Press, 2018).

Daniel Stoljar is Professor of Philosophy at the Australian National University in Canberra and a Fellow of the Australian Academy of the Humanities. He is the author of *Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness* (Oxford UP, 2006), *Physicalism* (Routledge, 2010), and *Philosophical Progress: In Defense of a Reasonable Optimism* (Oxford UP, 2017).

Galen Strawson holds the President's Chair of Philosophy at the University of Texas at Austin. His books include *Freedom and Belief* (Oxford UP. 1986), *The Secret Connexion* (Oxford UP, 1989), *Mental Reality* (MIT Press, 1994), *Selves* (Oxford UP, 2009), *Locke on Personal Identity* (Princeton UP, 2011), and *The Subject of Experience* (Oxford UP, 2017).

Donovan Wishon is Associate Professor of Philosophy at the University of Mississippi. He is the author of a number of papers on Bertrand Russell's theories of mind, matter, language, and knowledge, and his coedited volume *Acquaintance, Knowledge, and Logic* (with Bernard Linsky) received the 2016 Bertrand Russell Society book prize.

PREFACE TO THE HANDBOOK

Roughly speaking, panpsychism is the view that mentality is a fundamental feature of the world which exists ubiquitously throughout the world. The relevant kind of mentality is the most puzzling and strangest phenomenon in the entire universe: consciousness. The fundamental form of consciousness posited by panpsychism is presumably very remote from that of human consciousness and possessed of a primitive and unutterably simple nature. At a very basic level, the world is awake.

Mostly people intuitively find panpsychism implausible, but sometimes for bad reasons. For example, panpsychism does not imply that everything is conscious because not everything is a fundamental entity. Most things in the world are composite beings. Whether a particular composite being is conscious or not depends on the 'mental chemistry' which binds or transforms the fundamental consciousness of its constituents into more complex forms of consciousness. Some things, like us or our brains, have the right chemistry. Rocks do not. One of the major philosophical problems facing panpsychism – the so-called 'combination problem' – is to understand how mental chemistry works, or even to understand how it is so much as possible in the first place.

Still, panpsychism remains implausible to most. Yet it has experienced a remarkable renaissance of interest over the last quarter century. The reason for the rebirth of serious engagement is the intractable problem of consciousness. Despite immense progress in understanding the brain and its correlational relation to states of consciousness, we still really have no idea how consciousness could emerge from physical states and processes which are *presumed* to be entirely nonconscious in their fundamental nature. As Thomas Huxley famously quipped in 1876 'how it is that anything so remarkable as a state of consciousness comes about as a result of irritating nervous tissue, is just as unaccountable as the appearance of the Djinn, when Aladdin rubbed his lamp'. But somehow the physical world performs this feat on a regular basis. This suggests that consciousness is somehow a 'way of being physical'. My current experience is telling me something – something quite extraordinary – about the nature of the physical being I am. In the absence of belief in magical eruptions of metaphysically discontinuous properties, an almost inevitable line of thought drives us toward panpsychism.

Thinking about panpsychism brings together a host of core issues in the philosophy of mind, the philosophy of science, metaphysics, and even ethics. This book aims to introduce

and explore such issues from a wide diversity of viewpoints.

I give a broad philosophical introduction to panpsychism and its place in current debates about consciousness and its place in nature in the initial chapter of this book, 'A Panpsychist Manifesto'.

Here I will very briefly introduce the handbook itself. The goal of the handbook is to provide a high-level comprehensive examination and assessment of panpsychism – its history and contemporary development. It will be useful both to students and scholars as both an introduction and cutting-edge philosophical engagement. For anyone with a lively intellect interested in a philosophical approach to things, the handbook should supply fascinating and enlightening reading. As noted, the topics covered are highly diverse, representing a spectrum of views on the nature of mind from various standpoints which take panpsychism seriously.

The handbook is divided into 28 chapters roughly sorted into a number of themes. The first is a set of historical reflections ranging from Ancient Greek philosophy, Buddhist thought, Early Modern philosophy to the rich period for panpsychism of the 19th and early 20th centuries.

The second section deals broadly with various forms of panpsychism. These encompass distinct kinds of consciousness which might form the basic building block of a panpsychic view of the universe as well as different ways that panpsychism might describe the world, notably in the distinction between micropsychism (which assigns primitive mental aspects to elementary physical entities) to cosmopsychism (which takes the entire universe to be the fundamental conscious entity).

Next, in the third section, panpsychism faces up to a number of critical alternatives which are distinctive for taking panpsychism seriously and confronting their preferred doctrine directly with it. These include less well-known approaches such as subjective physicalism and cognitive pluralism, as well as the better known alternatives of neutral monism and the modern rebirth of Russellian monism.

In the final section, the chapters provide a variety of viewpoints on how panpsychism would or could actually work or be articulated in a more precise form. These chapters consider, for example, the question of how or whether individual subjects of experience could be 'summed' or 'combined' into more complex subjects, whether quantum mechanics can help fund a panpsychic outlook, whether causation itself can underpin some form of panpsychism and the core question of the real nature of the infamous combination problem.

The handbook also contains two 'keynote' articles. One is by Galen Strawson, who has long been one of the most forceful advocates for the intelligibility and possible intellectual benefits of panpsychism. He provides a close examination of the relation between panpsychism and our understanding of the physical, *the* issue which is at the heart of panpyschism. The second is by David Chalmers, who played a large role in the rebirth of interest in panpsychism at the end of the 20th century with his famous invocation of the 'hard problem' of consciousness and has since done much to deepen our understanding of panpsychism. His paper examines a comprehensive set of forms of idealism, the connection with panpsychism, and considers the question whether idealism is the natural endpoint beyond panpsychism for those moved by the seeming intractability of the problem of

consciousness.

I would like to thank all the contributors to the handbook who took my editorial interventions in good spirit and waited patiently for the publication to come together. I also thank the team at Routledge who have shown understanding and patience across a long gestation period, and who were always fully supportive, not least – to my mind – in tracking down the beautiful Emily Carr painting which graces the handbook's cover.

1
INTRODUCTION

A Panpsychist Manifesto

William Seager

The world is awake. That can stand as a slogan for *panpsychism*: the view that I will understand here as holding that consciousness is fundamental and ubiquitous in nature. This does not mean that everything is conscious. Whether a particular non-fundamental entity is conscious will depend upon the arrangement of its fundamental constituents given some presumed laws of 'mental chemistry'[1] which govern the emergence of complex forms of consciousness. So in bare outline panpsychism presents a familiar picture of fundamental features interacting in ways to generate more complex forms, it's just that the catalog of the fundamental includes consciousness. Nor does panpsychism entail that sophisticated, high-level, human-like consciousness is ubiquitous. The term 'consciousness' is notoriously hard to define and the victim of multitudes of more or less well motivated (re)definitions. I aim for a minimal conception. For contrast, compare this expansive notion of consciousness, plucked merely for illustrative purposes from Aaronson (2016): 'displaying intelligent behavior (by passing the Turing Test or some other means) might be thought a necessary condition for consciousness'. On the minimal conception, consciousness does not at all require that ability to pass the Turing test. Feeling pain (or any other sensation) alone is sufficient for consciousness; consciousness implies only sentience. It's worth noting this because there is a somewhat pernicious ambiguity lurking here, that between a property and the evidence we have for ascribing it. Although still inaccurate, Aaronson's dictum is closer to the truth if we change the final phrase to 'a necessary condition for the ascription of consciousness'. But note that we can have *theoretical* reasons for ascribing a property without there being any direct observational evidence for the ascription. So, the kind of minimal consciousness in question is not 'self-consciousness' or 'transcendental subjectivity', or awareness of the self as a subject, or awareness of one's own mental states, or the ability to conceptualize one's own mental states as such. Consciousness is simply sentience, or the way things are present (to the mind).

It is undeniable that panpsychism is intuitively implausible. It is frequently subject to derision by philosophers, being labeled 'absurd' (Searle 2013) and 'ludicrous' (McGinn 1999: 97). Even sympathizers have qualms. Thomas Nagel worries that panpsychism carries the taint of 'the faintly sickening odor of something put together in the metaphysical laboratory' (1986: 49). Such denigrations stem from a certain confidence – misplaced I think – in our pre-existing conception of the nature of the physical world and a rather strange lack of confidence in our conception of subjective consciousness. We think we know what matter is, and we think we thereby know that it just is *not* the kind of thing which is or could be intrinsically conscious. In fact, though commonplace the former belief is demonstratively mistaken, which leaves open the status of the latter claim.

This is not a new thought. In one form or another it dates back at least to 19th-century writers such as Ernst Mach and William James, and in the early 20th century to Bertrand Russell and Arthur Eddington. The basic idea that the nature of matter is not obvious just from our daily interactions with material objects forms the core of Noam Chomsky's intriguing but somewhat obscure views on the mind-body problem, seemingly leading to Chomsky requiring/expecting a physical-science revolution before consciousness can be understood as a natural phenomenon (see 2000: ch. 4). The staunch anti-panpsychist John Searle sometimes also suggests the need for a revolution in science in order to understand how consciousness is a 'biological phenomenon', likening the situation to that of physics prior to the introduction of the electromagnetic field (1992: 101 ff.). Searle's claim that there is no special problem of consciousness because of the 'biological powers' of the brain has always struck me as, therefore, rather strange and one wishes he would explain, not in the details that can be left to science, but just the general mechanism of emergence by which the biological generates consciousness which would then reveal to us where to look for the neurological details. There is considerable recent work devoted to understanding the nature of the physical (see e.g. Montero 2009; Wilson 2006) and much current interest in the associated doctrine of Russellian Monism (see e.g. Alter and Nagasawa 2015). This idea is forcefully expressed in defense of panpsychism in the work of Galen Strawson (see e.g. 2006, 2003 and Strawon's chapter in this volume).

In the face of this natural antipathy, consideration of the philosophical advantages of panpsychism is the best way for it to gain sympathy. The primary motivator for panpsychism is the problem of consciousness. Over the last fifty years we have witnessed staggering advances in our knowledge of the brain and our ability to observe it in action. As we enter into the search for the elusive neural correlates of consciousness there are truly remarkable developments in mapping the relations between states of consciousness and neural activity.

Perhaps the most astonishing example of our growing knowledge, which is now underpinning practical clinical interventions, is the work of Adrian Owen, who investigates people diagnosed as being in a profound vegetative state. Such patients are incapable of making any overt behavioral response which could signal residual consciousness and have been regarded as completely unconscious. Owen has found that a disturbing number of such patients are in fact 'locked in' – fully conscious but cut off completely from their bodies. Using real-time functional MRI, Owen is able to interact with supposedly vegetative

subjects by asking them to imagine various activities, such as playing tennis, or walking around their home. It is possible to identify the (no doubt partial) neural correlates of what the subject is imagining and thus open up a channel of communication. Owen is currently working on non-MRI solutions which patients will be able to use outside his laboratory, perhaps even at home.[2] This is what we can do now. There is little doubt that our ability to discover and track neural correlates of conscious states will expand enormously over the next decades.

Nonetheless, the infamous 'explanatory gap' (see Levine 1983) between the physical states of the brain and consciousness remains. Finding neural correlates does not show us how the brain manages to generate or realize consciousness and thereby solve David Chalmers's 'hard problem' (see his 1995, 1996). It seems likely that we will in the near future discover deeper and much more specific neural signatures of conscious states, perhaps involving neural synchronization or distinctive neural or subneuron dynamical activity associated with consciousness. That too won't reveal how consciousness is produced by the brain: these mechanisms are only more fine-grained correlations. There is a problem of principle here.

Why is that? As hinted previously, the core of the problem is the apparent mismatch between the nature of the physical world as we understand our fundamental theories to have revealed and the subjective, 'what it is like' aspect of minimal conscious experience. It feels like something to be awake and this just seems utterly foreign to how we regard and, for many, how we *ought* to regard the material world.

I take it that physics provides our scientific understanding of that world at the most fundamental level. However, its theories are perpetually provisional and the more or less background metaphysical pictures which they both suggest and spring from have a distressing history of being radically overturned by those of newer theories. And yet it is undeniable that we have accumulated knowledge. This is possible because the metaphysical picture which lurks within or behind our physical theories is not essential to their use and need not be preserved across scientific revolutions, even as explanatory and predictive power is retained and expanded. What is preserved is the relational or structural systems which physical theory maps out and by which it is confirmed and becomes predictively successful.[3] These structures are all, ultimately, relations between observable quantities for which we have labels such as 'mass', 'electric charge', 'momentum', etc., all definable in terms of observable *motion*. Over centuries of development, physical theory has become successively more complex (for example, many kinds of charge beyond that of electric are now recognized) but all new hypotheses link to the relations between observable quantities, albeit sometimes indirectly.

It is unsurprising that nowhere in this system do we find subjectivity, nor in the development of physics is there any need to posit a subjective aspect to nature. The explanatory gap is exactly the problem of how a world which is supposed to be completely described at the fundamental level by a science which has no place or need for subjectivity nonetheless somehow includes the subjective aspect of the world we call 'consciousness'.

The problem has been recognized for a very long time. In 1714 Leibniz expressed it with his 'mill argument' in the *Monadology*:

> Imagine there were a machine which *by its structure* produced thought, feeling, and perception. We can imagine it as being enlarged while maintaining the same relative proportions, to the point where we could go inside it, as we would go into a mill. But if that were so, when we went in we would find nothing but pieces which push one against another, and never anything to account for a perception … perception, and everything that depends on it, is inexplicable by mechanical principles.
>
> (1714/1989: 215, my emphasis)

Note that Leibniz makes the anti-structuralist point that the causal organization of the mill, or the brain, cannot provide an explanation of the appearance of consciousness even if it is correlated with it.[4]

Leibniz was targeting the so-called 'mechanical philosophy' which, roughly speaking, asserted that the material world was, as Newton put it, such that 'God in the Beginning form'd Matter in solid, massy, hard, impenetrable Particles, of such Sizes and Figures, and with such other Properties, and in such Proportion to Space, as most conduced to the End for which he form'd them' (1730/1979, Query 31). Of course, Newton did not stop with this characterization but left the strict mechanical view behind by adding that 'these Particles have not only a Vis inertiæ, accompanied with such passive Laws of Motion as naturally result from that Force, but also that they are moved by certain active Principles'. Gravity is but one example of such 'principles', which he also called 'Powers, Virtues, or Forces'.[5] Newton expected a scientific chemistry to emerge in time based upon them.

A philosophically purified version of mechanism was much later articulated by C. D. Broad which can stand as the mechanistic ideal, whose essence is:

> (a) a single kind of stuff, all of whose parts are exactly alike except for differences of position and motion;
> (b) a single fundamental kind of change, viz, change of position …
> (c) a single elementary causal law, according to which particles influence each other by pairs …
> (d) a single and simple principle of composition, according to which the behavior of any aggregate of particles, or the influence of any one aggregate on any other, follows in a uniform way from the mutual influences of the constituent particles taken by pairs.
>
> (1925: 44–5)

This vision is so ethereal that nothing like it has ever been seriously entertained. But the actual mechanical worldview and its successors for a couple of centuries of furious and spectacularly successful development can be seen as implicitly inspired by this pure vision. It expresses well a picture of the world – let's call it 'LEGO® world' – formed of a very large number of very small parts which are metaphysically independent of each other, have individual identities (albeit ones of very little interest) but which can interact by local causation. The familiarity we have with things like this, think marbles or, indeed, LEGO bricks, is what funds confidence in our conception of matter or the physical and makes it seem intuitive and almost obvious … and very distant from subjectivity.

The mechanical world view leveraged this picture into an intuitive positive conception of the nature of matter: it came in chunks akin to the small particles we are familiar with in ordinary experience: impenetrable, capable of motion and – thanks ultimately to God's decree – observant of the fundamental rules or laws of nature which governed how these pieces interacted (e.g. conservation of energy). Exactly how and why these material units are 'forced' to obey the laws of nature remained (and remains) somewhat obscure. Perhaps the laws' power is just a primitive metaphysical fact which links properties with the appropriate level of modal force. Perhaps the laws somehow follow from the causal powers of the fundamental entities of the world. Maybe the laws are a mere catalog of universal regularities, or meta-regularities across a set of possible worlds (the 'nomologically possible worlds'). Perhaps the laws are an imposition of the conscious mind which 'imposes' them upon an intrinsically chaotic universe (we find laws of nature because we could not exist in those regions of possibility that were 'too' chaotic; a kind of anthropic Kantianism).

In any case, this conception of matter excluded consciousness as one of its properties (except perhaps if God directly and miraculously 'superadded' it to a material system[6]) and, in any case, there was no need to posit consciousness to generate all the forms and activities in which matter could participate. As to these forms, we now know that LEGO world is capable of implementing a Turing machine so such a world could at least simulate anything which can be computed (ignoring speculative hyper-computational devices), including quantum mechanics. Such a simulation might be very slow but one should not confuse the simulation runtime with the time internal to the simulation. Perhaps, so to speak, it might take seconds of simulation time to compute one yoctosecond of internal simulation time, but we can suppose that the simulation lives in an eternal and spatially infinite world so there is no principled limitation here (we can suppose the physics of the simulation world is unconstrained by anything except its being a LEGO world). Presumably, our entire universe, or at least the observable universe, can thus be simulated given that the number of possible initial conditions (the degrees of freedom of the big bang) is finite (which seems to be implied by current theory, e.g. the holographic principle and possible theories of quantum gravity).

So the LEGO world conception is in fact a spectacularly powerful one and it is based upon an intuitively attractive conception of the nature of matter. One reason panpsychism seems weird to people is that they have implicitly absorbed something like the mechanistic view of the material world and its conception of the nature of the physical. But notwithstanding the power of the LEGO world picture, its positive conception of the physical was exploded in the first half of the 20th century. Matter does not form the 'LEGO world' imagined by the early mechanists. It turns out that matter is nothing at all like 'matter' was supposed to be.

Of course, it is quantum mechanics that has revealed this to us. This lesson is perhaps the number one thing that quantum mechanics is 'trying to tell us'. Quantum entanglement seems to hint that nature is holistic and that a world of independent material parts is merely a kind of emergent approximation. This has been noted many times by physicists beginning almost from the birth of quantum mechanics:

- Erwin Schrödinger: a particle certainly is … not a durable little thing with individuality (1952: 241).
- Hans Primas: the idea … that the material world is … structured by some kind of interacting 'elementary systems' is in sharp contradiction [with] quantum mechanics (1998: 88).
- Basil Hiley: quantum phenomena require us to think in a radical new way, a way in which we will have to ultimately give up both the notion of particles and fields (1999: 116).
- David Bohm: the entire universe must, on a very accurate level, be regarded as a single indivisible unit in which separate parts appear as idealizations permissible only on a classical level of accuracy of description (1951: 167).

Our best quantum theory asserts that fundamental reality is composed not of material particles at all but rather strange universal fields, the temporary excitations of which can appear to our experiments as particle-like apparitions. These fields appear to exist in an extremely (infinite?) high-dimensional space which cannot be directly identified with the space and time of experience.

But despite the radical revision required by the quantum mechanical picture of the world, our new physics has no need to add *consciousness* to the quantum fields to generate the physical phenomena and structure that we can observe.[7] The structures and systems of relations amongst the physical attributes of the world are generated from the purely physical fundamental features posited by physics with no hint of intrinsic subjectivity. So the great revolution in physics occasioned by the discovery of quantum mechanics has not by itself closed the explanatory gap and solved the hard problem of consciousness. In fact, as noted, some aspects of quantum physics, most famously the nature of observation or measurement, may suggest a fundamental role for consciousness[8] but, even if it is real, it is a role which figures *outside* of the system described by physical theory, breaking the law which governs the evolution of the quantum state.

The end of LEGO world and the mechanistic, or neo-mechanistic, picture of the world means we have lost any viable positive conception of the fundamental nature of matter. What we are left with is nothing more than the relational structure linked to observables, which physics maps out for us with such stunning success. It is thus also the case that there is no direct impediment to the panpsychist hypothesis stemming from an acceptable positive conception of the nature of the physical.

Lacking an explanation of consciousness in physical terms and lacking any conception of material reality beyond the structural, the panpsychist steps into the opening and suggests that perhaps the fundamental reality of the physical world itself partakes of some aspect of subjective consciousness. Again, this does not mean that quantum field particle states are reflecting about their own existence, enjoying a rich inner life akin to our own (though some vast multi-particle states are, namely ourselves). But minimal, unsophisticated and unreflective consciousness is much more common than its opposite, as witnessed by the host of pretty clearly conscious animals, some of which must have extremely limited forms of awareness. All the panpsychist needs to posit is that some form of subjectivity, some kind

of primitive feeling, is at the foundation of the physical world.

No positive conception of the physical precludes this posit, since we have no such conception. The hypotheses of physics, of course, make no mention of consciousness but this is no surprise given the structuralist constraints of physical science. But, roughly speaking, something beyond structure is needed to make reality concrete. This point was perhaps made most precisely and forcefully by Max Newman (1928) in his critique of Russell's attempt at a purely structural account of science. Given any system of entities, a merely abstract specification of structure as, say, sets of ordered n-tuples, is revealed to be already present (sets exist if their members do). Something has to 'select' a certain structure as 'real' or concrete and, obviously, this 'something' has to go beyond simply additional relational structure. This point should not be over interpreted. It does not imply the return of objects in the sense of independent units with persistent and genuine individuality. It is simply the requirement that some form of concrete reality is needed to, so to speak, realize what are otherwise pure abstractions.[9] Why not, asks the panpsychist, let the one nonstructural reality we are already acquainted with and for which we do have a positive conception, namely our own subjectivity, stand as this foundation?

I will concede that this sort of consideration does not make panpsychism exactly plausible. It strikes many as beyond strange to think of the world as awake and kind of 'humming' with a primitive consciousness suffused throughout it as if there was an extra dimension of basic subjectivity. Many or most would still prefer a physicalist metaphysics. But it seems to me that, at bottom, there are only two alternatives which abide with standard physicalism.[10]

The first is emergentism. Like the word 'consciousness', the term 'emergence' is fraught with ambiguity. The minimal sense of the notion is simply that of a system having properties which are not possessed by its parts. Familiar and uncontroversial examples abound: the liquidity of water, the vortex structure of a tornado or hurricane. Commonplace but more exotic and controversial forms are things like the collective intentionality of groups or nations. In any case, the world is awash in emergence: almost everything you could think of has properties which are not shared by the fundamental physical entities postulated by physics. But this sort of standard or conservative emergence based upon complexity does not engender an explanatory gap. When vast numbers of hydrogen and oxygen atoms, for example, get together just right (and noting that these atoms, and their constituent nucleons are themselves emergent entities of the same sort) we get liquid water, which is a remarkable substance whose complexities are still not fully understood, although we do more every day with 'ab initio molecular dynamics simulations' of critical temperatures, density changes and the like (see e.g. Morawietz et al. 2016). Such efforts reveal that there is no doubt about the general route from oxygen and hydrogen to the properties of water. From one point of view, this should not be at all surprising. The properties of water at issue are themselves purely structural and a natural target for the structure-based systems explored by more fundamental physics.

On a much grander level, we have learned a tremendous amount about the emergence of the classical world from its quantum underpinnings via decoherence (see the canonical text Joos et al. 2003) and the rise of effective field theory, understood via the renormalization

group procedure, which reveals why the world is amenable to our theoretical descriptions even if we can be quite sure that these descriptions are in themselves incomplete and ultimately inaccurate. Thus quantum field theory itself is regarded as a merely empirically adequate emergent within a delimited sector of nature: 'we have learned in recent years to think of our successful quantum field theories ... as "effective field theories", low-energy approximations to a deeper theory that might not even be a field theory' (Weinberg 1995: xxi). Note that from the structuralist viewpoint the prospect of potentially radical theoretical transformation is not a big problem. The predictive and explanatory successes of field theories remain since the structure they revealed remains after theory change.

Let's group this widespread and internally quite diverse form of emergence under the term of 'conservative emergence'. Its twin hallmarks are first, that there is an intelligible route from the constituents of a system, broadly conceived to include everything from atoms to field states to theoretical structures, to its emergent properties, even if complexity will often preclude precise predictions of exact values of the emergent features; and second, that the emergent features will be largely describable and explicable in terms of their own proprietary laws, albeit always subject to potential interference from the more fundamental submergent underpinnings. It is an amazing fact about our world that it is structured by a hierarchy of relatively autonomous emergent domains (of course, if it were not, we would not be here).

The emergence of subjectivity presents a problem of an entirely different order and seems to call for more than conservative emergence. Ever more complex calculations of the behavior of ever more complicated physical systems will not get us to consciousness, even if it should prove possible to model the behavior of human beings, say via sophisticated computational neural simulations. There is no intelligible route from the relational structures found in our physical science to the intrinsic property of subjectivity. We are thus left with a *radical* emergence in which subjectivity simply appears at some, seemingly arbitrary, point in the physical development of the universe rather like Locke's fantasy of God directly 'superadding' consciousness to matter. Although logically possible, such a radical emergence is entirely at odds with the whole point of the physicalist enterprise. Furthermore, there is no reason to think that a radically emergent feature is itself physical in nature; quite the opposite in fact, given that all the physical features we know are either fundamental or conservative emergents that arise from assemblages and interactions of, ultimately, the fundamental features of the physical world.

Absent emergence, the second alternative is denial. Perhaps there is nothing more to the world than the relational structures posited by physics, along with all the conservatively emergent outgrowths of basic physical processes. There are two immediate problems with such a heroic response. The first is that we are left in ignorance about the underlying nature of matter. I suppose the best physicalist reply is that this is simply an unavoidable and fundamental limitation to our knowledge, regrettable perhaps but not such as to recommend endorsing panpsychism over epistemic humility (see Langton 1998; Lewis 2009).

The second problem seems much more serious: denying the reality of subjective experience is a desperate measure and one of dubious coherence. The immediately most obvious argument for the incoherence of the position is very simple:

1. If consciousness is an illusion then it merely seems that it exists.
2. But if anything *seems* to exist, that seeming is a state of consciousness.
3. Therefore consciousness (states of consciousness) exists.

The most developed ultra-eliminativism about experience itself is that of Daniel Dennett which attempts, cleverly but ultimately unsuccessfully, to work around this argument. Dennett proposes that we suffer an illusion that subjective experience exists but that this illusion is a purely cognitive 'thought-illusion' which can be explained in terms of a science of entirely nonconscious intentional states (see e.g. Dennett 1991; for criticism see Seager 2017). Dennett exploits the fact that the term 'seems' has both an experiential and epistemic sense (contrast 'the sun seems to be rising' versus 'it seems that Trump will irreparably harm America'). The latter sense does not immediately demand any experiential component, which is why Dennett can say without falling into immediate contradiction that '[t]here seems to be phenomenology… . But it does not follow from this undeniable, universally attested fact that there really is phenomenology' (1991: 366).

Of course, this line of attack requires some explication of a purely cognitive, non-experiential theory of thought content which Dennett has attempted to supply with his theory of intentional stances. A particularly sophisticated deployment of this account leads to what he calls the method of heterophenomenology. Applying it to those misguided enough to believe in consciousness, heterophenomenology

> involves extracting and purifying texts from (apparently) speaking subjects, and using those texts to generate a theorist's fiction, the subject's heterophenomenological world. This fictional world is populated with all the images, events, sounds, smells, hunches, presentiments, and feelings that the subject (apparently) sincerely believes to exist in his or her (or its) stream of consciousness.
>
> (Dennett 1991: 98)

The basic problem with this approach, as I see it, is that it fairly obviously has a tacit dependence on the experience and consciousness of the notional 'researcher' who is performing the heterophenomenological exercise.

One way to see this is to take an epistemological route. Following the old path of Descartes, let us each ask ourselves whether we know that something exists, or that something is happening right now. It is evident that we have access to a realm of data which by itself guarantees that yes, indeed, something exists. But the relational structures of physics cannot provide that guarantee since their existence is not a certainty: what is certain is the realm of subjective experience. If we posit that this realm does not exist, then we are left with the disturbing possibility that, given how things are now, nothing at all exists. If anything is clear it is that that thought is self refuting as soon as it is consciously thought. We do have access to a realm of immediately available and unassailable self-knowledge, albeit of a quite limited form (which limitations Descartes also emphasized). Denying this is to undermine the claim that we can know we are thinking at all, or indeed, that we know that anything exists. It is much more plausible that consciousness exists than that we cannot unassailably assert certain propositions which we understand.

Granting that conscious experience exists (as if this were a concession!) does not entail

panpsychism. We can retain an epistemic humility about the underlying nature of the material world along with the – in this context now somewhat puzzlingly assertive – denial that it involves subjectivity. And we can always retain the hope that a conservatively emergent path from neuroscience (or elsewhere in the sciences) will illuminate the generation of consciousness by a world bereft of fundamental subjectivity. But the former view is not mandatory and there is not the slightest hint of how the latter might be accomplished.

Panpsychism promises to integrate our scientific and 'personal' view of the world and do so in a way that respects both the completeness of the physical picture of the causal structure of the world it investigates and the role of consciousness itself. The price to pay is admission of subjectivity into the foundation of the world as one of its fundamental features.

Needless to say, panpsychism faces many problems beyond its initial air of implausibility. But these are problems which can be philosophically explored, analyzed and perhaps solved. For example, one might wonder why, if panpsychism is true, we see no sign of subjectivity at the fundamental physical level (electrons do not seem to be experiential subjects). But in fact the only direct revelation of subjectivity is in our own experience and the panpsychist interprets this as evidence, that 'consciousness … provides a kind of "window" on to our brains' thereby revealing 'some at least of the intrinsic qualities of the states and processes which go to make up the material world' (Lockwood 1989: 159). The fact that very simple physical systems do not exhibit complex behavior should not be surprising but it's hardly clear that this *shows* they lack a subjective aspect.

The most vexing difficulty facing panpsychism is the so-called 'combination problem' (see James 1890/1950: ch. 6; Seager 1995), which is the issue of how the primitive foundational aspect of subjectivity postulated by the panpsychist builds itself into the sophisticated and rich forms of consciousness which creatures like us enjoy. This is not the place to discuss this in detail. Suffice to say that recent work has offered a variety of possible approaches to the combination problem (see many of the papers in Brüntrup and Jaskolla 2016, especially Chalmers 2016 and Goff 2016, and many of the chapters in this volume). Some of these approaches use the reasonably familiar relation of co-consciousness, some use a relation of phenomenal blending, some devise a notion of phenomenal bonding and some develop the idea that primitive experiential aspects 'fuse' into new forms, superceding the originals. It is also possible to re-orient the problem by adopting a radical holism, in which the single fundamental entity is the universe as a whole. Panpsychism then becomes what is called 'cosmopsychism', and the combination problem becomes the 'de-combination' problem. This somewhat Spinozistic view is as yet little explored but see Goff (2017) as well as the chapters of Albahari and Goff in this volume. All of these ideas face serious objections, but they all stand as potential answers to the combination problem. The point here is that the combination problem is one that can be addressed fruitfully. It is not a showstopper.

Philosophical exploration and development of the panpsychism is possible and has already proved fruitful. Panpsychism is a viable solution to the traditional mind-body problem and addresses the more modern specific problem of consciousness. We can throw

off the shackles of an outmoded and falsely restrictive conception of the physical and declare that the world is awake.

Notes

1. This term goes back to John Stuart Mill (see 1843/1963: 108–9); see also Nagel (1979).
2. Owen has written a beautiful and fascinating popular exposition of his work entitled *Into the Gray Zone* (2017); a scholarly presentation of methods and results can be found in Owen (2008).
3. For philosophical discussion of scientific structuralism, the doctrine that science provides only knowledge of structural or relational features of reality see French (2014).
4. Leibniz's celebrated solution to the mind-body problem, the pre-established harmony (1695/1997) between the material and mental realms, *predicts* the existence of neural correlates of consciousness. Leibniz (1696/1997) provides an interesting discussion of the three possible relations between matter and consciousness: causal, miraculous and harmony.
5. Newton's acceptance of such perhaps 'occult' powers was anathema to strict mechanists who denied both action at a distance and hoped for a physics based solely upon interparticle collision. Leibniz himself, along with several other luminaries, attempted to develop vortex-based accounts of planetary motion which avoided Newtonian gravitation with some incomplete success (see Aiton 1972).
6. Something Locke famously asserted was within the power of God and to which Leibniz agreed, though according to him it could be accomplished only via an objectionable 'perpetual miracle' (see 1704/1996: 67).
7. It may need to add consciousness to the picture in order to get determinate observations, but that is an extra feature that has always seemed unscientific. Very few, but not zero, physicists believe that the solution to the measurement problem essentially involves consciousness (see McQueen (2019) for a popular level philosophical discussion).
8. The specific idea that measurement requires the intervention of a conscious observer appears in somewhat veiled ways in Bohr and von Neumann. It is explicitly advanced in London and Bauer (1939/1983) and Wigner (1962). A modern proponent is Henry Stapp (e.g. 1993).
9. For discussion of the subtleties of Newman's objection to Russell see Demopoulos and Friedman (1985); see also van Fraassen (2007) and the radical pro-stucturalist discussion in Ladyman et al. (2007).
10. There are non-physicalist alternatives to panpsychism. Idealism, for example, retains defenders and has sparked some renewed interest in recent philosophy (see e.g. Foster 2008; Pelczar 2015; Chalmers's paper in this volume). Various forms of dualism are further alternatives that still have a few defenders (see e.g. Robinson 2004; Lavazza and Robinson 2014). Arguing for panpsychism over such non-physicalist views cannot be undertaken here, but roughly speaking it is panpsychism's ability to integrate with the scientific view of the world and provide a solution to the problem of mental causation that are its main advantages.

References

Aaronson, Scott (2016). 'The Ghost in the Quantum Turing Machine'. In S. Barry Cooper and Andrew Hodges (eds.), *The Once and Future Turing*. Cambridge: Cambridge University Press, pp. 193–296.

Aiton, Eric (1972). *The Vortex Theory of Planetary Motions*. London: Macdonald.

Alter, T., and Nagasawa, Y. (eds.) (2015). *Consciousness in the Physical World: Perspectives on Russellian Monism*. Oxford: Oxford University Press.

Bohm, David (1951). *Quantum Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Broad, C. D. (1925). *Mind and Its Place in Nature*. London: Routledge and Kegan Paul.

Brüntrup, G., and L. Jaskolla (eds.) (2016). *Panpsychism*. Oxford: Oxford University Press.

Chalmers, David (1995). 'Facing Up to the Problem of Consciousness'. *Journal of Consciousness Studies*, 2 (3): 200–19. (Reprinted in J. Shear (ed.), *Explaining Consciousness*. Cambridge, MA: MIT Press, pp. 9–32, 1997).

Chalmers, David (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.

Chalmers, David (2016). 'The Combination Problem for Panpsychism'. In G. Brüntrup and L. Jaskolla (eds.), *Panpsychism*. Oxford: Oxford University Press, pp. 229–48.

Chomsky, Noam (2000). *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.

Demopoulos, William, and Friedman, Michael (1985). 'Bertrand Russell's the Analysis of Matter: Its Historical Context and Contemporary Interest'. *Philosophy of Science*, 52 (4): 621–39.

Dennett, Daniel (1991). *Consciousness Explained*. Boston: Little, Brown & Co.

Foster, John (2008). *A World for Us: The Case for Phenomenalistic Idealism*. Oxford: Oxford University Press.

French, Steven (2014). *The Structure of the World: Metaphysics and Representation*. Oxford: Oxford University Press.

Goff, Philip (2016). 'The Phenomenal Bonding Solution to the Combination Problem'. In G. Brüntrup and L. Jaskolla (eds.), *Panpsychism*. Oxford: Oxford University Press, pp. 283–304.

Goff, Philip (2017). *Consciousness and Fundamental Reality*. Oxford: Oxford University Press.

Hiley, Basil (1999). 'Active Information and Teleportation'. In D. Greenberger, W. Reiter, and A. Zeilinger (eds.), *Epistemological and Experimental Perspectives on Quantum Physics*. Dordrecht: Kluwer.

James, William (1890/1950). *The Principles of Psychology*, vol. 1. New York: Henry Holt and Co. (Reprinted in 1950, New York: Dover, ed.X. (Page references to the Dover edition.)).

Joos, Erich, Dieter Zeh, H. et al. (2003). *Decoherence and the Appearance of a Classical World in Quantum Theory*. Berlin: Springer-Verla.

Ladyman, James, Ross, Don et al. (2007). *Everything Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press.

Langton, Rae (1998). *Kantian Humility: Our Ignorance of Things in Themselves*. Oxford: Oxford University Press.

Lavazza, A., and Robinson, H. (eds.) (2014). *Contemporary Dualism: A Defense*. London: Routledge.

Leibniz, Gottfried Wilhelm (1695/1997). 'New System of the Nature of Substances and Their Communication, and of the Union Which Exists Between the Soul and the Body'. In R. Woolhouse and R. Francks (eds.), *Leibniz's 'New System' and Associated Contemporary Texts*. Oxford: Oxford University Press, pp. 7–36.

Leibniz, Gottfried Wilhelm (1696/1997). 'Letter to Basnage'. In R. Woolhouse and R. Francks (eds.), *Leibniz's 'New System': And Associated Contemporary Texts*. Oxford: Oxford University Press, Clarendon Press, pp. 62–4. (See C. I. Gerhardt (ed.) *Die philosophischen schriften von Gottfried Wilhelm Leibniz*, vol. 4. Berlin: Weidmann, 1880, p. 496ff).

Leibniz, Gottfried Wilhelm (1704/1996). *New Essays on Human Understanding*. Cambridge: Cambridge University Press. (J. Bennett and P. Remnant (trans. and ed.). Although this work was not published by Leibniz it was an essentially complete work, composed around 1704–5).

Leibniz, Gottfried Wilhelm (1714/1989). 'Monadology'. In Roger Ariew and Daniel Garber (eds.), *G. W. Leibniz: Philosophical Essays*. Indianapolis: Hackett, pp. 214–24.

Levine, Joseph (1983). 'Materialism and Qualia: The Explanatory Gap'. *Pacific Philosophical Quarterly*, 64: 354–61.

Lewis, David (2009). 'Ramseyan Humility'. In David Braddon-Michell and Robert Nola (eds.), *Conceptual Analysis and Philosophical Naturalism*. Cambridge, MA: MIT Press, Bradford Books, pp. 203–22.

Lockwood, Michael (1989). *Mind, Brain and the Quantum*. Oxford: Blackwell.

London, Fritz, and Bauer, Edmond (1939/1983). 'The Theory of Observation in Quantum Mechanics'. In J. Wheeler and W. Zurek (eds.), *Quantum Theory and Measurement*. Princeton: Princeton University Press, pp. 217–59. (Originally published as 'La théorie de lobservation en mécanique quantique'. In *Actualités scientifiques et industrielles*, no. 775, Paris: Heinemann, 1939.).

McGinn, Colin (1999). *The Mysterious Flame: Conscious Minds in a Material World*. New York: Basic Books.

McQueen, Kelvin (2019). 'Does Consciousness Cause Quantum Collapse?'. *Philosophy Now* URL=https://philosophynow.org/issues/121/Does_Consciousness_Cause_Quantum_Collapse.

Mill, John Stuart (1843/1963). *A System of Logic*, vol. 7–8 of *The Collected Works of John Stuart Mill*. Toronto: University of Toronto Press.

Montero, Barbara (2009). 'What Is the Physical?' In B. McLaughlin, A. Beckermann, and S. Walter (eds.), *The Oxford Handbook of Philosophy of Mind*. Oxford: Oxford University Press, pp. 173–88.

Morawietz, Tobias, Singraber, Andreas et al. (2016). 'How van der Waals Interactions Determine the Unique Properties of Water'. *Proceedings of the National Academy of Sciences*, 113 (30): pp. 8368–73. www.pnas.org/content/113/30/8368.abstract.

Nagel, Thomas (1979). 'Panpsychism'. In *Mortal Questions*. Cambridge: Cambridge University Press, pp. 181–95. (Reprinted in D. Clarke, *Panpsychism: Past and Recent Selected Readings*. Albany: SUNY Press, 2004.).

Nagel, Thomas (1986). *The View from Nowhere*. Oxford: Oxford University Press.
Newman, M. (1928). 'Mr. Russell's Causal Theory of Perception'. *Mind*, 37: 137–48.
Newton, Isaac (1730/1979). *Opticks, or, a Treatise of the Reflections, Refractions, Inflections and Colours of Light*. New York: Dover, ed.X.
Owen, Adrian M. (2008). 'Functional Neuroimaging of the Vegetative State'. *Nature Reviews Neuroscience*, 9 (3): 235–43.
Owen, Adrian M. (2017). *Into the Gray Zone: A Neuroscientist Explores the Border Between Life and Death*. New York: Scribner.
Pelczar, Michael (2015). *Sensorama: A Phenomenalist Analysis of Spacetime and Its Contents*. Oxford: Oxford University Press.
Primas, Hans (1998). 'Emergence in Exact Natural Science'. *Acta Polytechnica Scandinavica*, 91: 83–98.
Robinson, William (2004). *Understanding Phenomenal Consciousness*. Cambridge: Cambridge University Press.
Schrödinger, E. (1952). 'Are There Quantum Jumps? Part II'. *The British Journal for the Philosophy of Science*, 3 (11): 233–42.
Seager, William (1995). 'Consciousness, Information and Panpsychism'. *Journal of Consciousness Studies*, 2 (3): 272–88. (Reprinted in J. Shear (ed.), *Explaining Consciousness*. Cambridge, MA: MIT Press, 1997.).
Seager, William (2017). 'Could Consciousness Be an Illusion?' *Mind and Matter*, 15 (1): 7–28.
Searle, John (1992). *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.
Searle, John (2013). 'Can Information Theory Explain Consciousness'. *New York Review of Books*, January 10: 54–58.
Stapp, Henry (1993). 'A Quantum Theory of the Mind-Brain Interface'. In *Mind, Matter and Quantum Mechanics*. Berlin: Springer, pp. 145–72.
Strawson, Galen (2003). 'Real Materialism'. In L. Anthony and N. Hornstein (eds.), *Chomsky and His Critics*. Oxford: Blackwell, pp. 49–88. (Reprinted with new postscript, in T. Alter and Y. Nagasawa (eds.), *Consciousness in the Physical World: Perspectives on Russellian Monism*. Oxford: Oxford University Press, 2015, pp. 161–208.
Strawson, Galen (2006). 'Realistic Monism: Why Physicalism Entails Panpsychism'. *Journal of Consciousness Studies*, 13 (10–11): 3–31. (Reprinted in A. Freeman (ed.), *Consciousness and Its Place in Nature*. Exeter: Imprint Academic, 2006.).
van Fraassen, BasC. (2007). 'Structuralism(s) About Science: Some Common Problems'. *Proceedings of the Aristotelian Society Supplementary Volume*, 81 (1): 45–61.
Weinberg, S. (1995). *The Quantum Theory of Fields, vol. 1: Foundations*. Cambridge: Cambridge University Press.
Wigner, Eugene (1962). 'Remarks on the Mind-Body Problem'. In I. Good (ed.), *The Scientist Speculates*. London: Heinemann, pp. 284–302. (Reprinted in J. Wheeler and W. Zurek (eds.), *Quantum Theory and Measurement*. Princeton: Princeton University Press, 1983, pp. 168–81.).
Wilson, Jessica M. (2006). 'On Characterizing the Physical'. *Philosophical Studies*, 131 (1):

61–99.

PART I
Historical Reflections

2
PLATO AND PANPSYCHISM

Daniel Dombrowski

1. Introduction

David Ray Griffin's 1998 book *Unsnarling the World Knot: Consciousness, Freedom, and the Mind-Body Problem* alludes to a rope metaphor from Schopenhauer to the effect that the key philosophical problem since the 17th century (the mind-body problem) is, on prevailing assumptions, unsolvable. One can highlight several prominent philosophers of mind to illustrate why we have been tied in a 'world knot': William Seager has claimed that we have no idea whatsoever *how* consciousness 'emerges' from matter (1991: 195). Jaegwon Kim has held that we have reached a 'dead end' regarding the mind-body problem (1993: 367). Colin McGinn has alleged that we will *never* be able to understand the emergence of consciousness from the brain (1991: 1–2, 7). John Searle has suggested that most of mainstream philosophy of mind is 'obviously false' (1992: 3). And Galen Strawson has maintained that only a 'revolutionary' new way of thinking will enable us to respond adequately to the mind-body problem (1994: 92, 99). Although Daniel Dennett is a bit more optimistic regarding a solution to the mind-body problem on prevailing assumptions, even he has portrayed consciousness as a 'mystery' (1991: 21).

Some philosophers of mind think that panpsychism is the type of revolutionary thinking mentioned by Strawson that will enable us to unsnarl the world knot. In the context of the present chapter, however, Thomas Nagel's phrasing may be a bit more helpful when he says that 'radical' speculation is needed in order for such an unsnarling to be successful (1986: 10). I am taking the etymology of the word 'radical' seriously in that it comes from the Latin *radix* for 'root.' That is, contemporary panpsychists are offering an oxymoronic *radically new* approach to the mind-body problem as well as revitalizing an ancient solution to it. It is not often noticed in contemporary debates in philosophy of mind that panpsychism may very well be part of a *philosophia perennis* that goes back to Plato.

The spirit of the present chapter is captured well in a quotation from Josiah Royce, who once suggested that:

> Whenever I have most carefully revised my … standards, I am always able to see … that at best I have been finding out, in some new light, the true meaning that was latent in old traditions… . Revision does not mean mere destruction.
>
> (1908: 11)

When the questions are asked, how can experience arise out of, and act back upon, nonexperience?, or, how can consciousness arise out of, and act back upon, nonconsciousness?, the two major responses since the time of Descartes have been in terms of some form of dualism or some form of materialism. Indeed, the debate between dualists and materialists since the time of Descartes can be seen as entrenched.

But the situation is not as hopeless as it seems. This is due in part to the fact that some major philosophers of mind (e.g., Seager 1999; Strawson 2006) have moved toward panpsychism since the time when Griffin's book was published. The apparently entrenched character of the dualism–materialism debate should not prevent us from appreciating Plato's influence on Aristotle, Aristotle's influence on Leibniz, and the Leibnizian character of contemporary debates regarding panpsychism. For example, the whole point to Aristotelian and Thomistic hylomorphism was to suggest that matter without form and form without matter are the results of abstraction in that concrete reality is populated by 'formbodies' or 'mindbodies' or 'soulbodies,' to coin terms that try to capture the nondualistic character of hylomorphs. The Aristotelian idea that mind is, in a manner, all things is a protest against the view that mind is what is left over when one abstracts away from either behavior or matter. Rather, matter in motion just *is* mind in some fashion, hence, on one plausible reading of Aristotle, panpsychism is implicitly affirmed.

Leibniz is explicitly a panpsychist in his *Monadology*, but Leibniz's similarity to Aristotle is not often noticed. Newton insisted that an active principle had to be operative *in* nature, and not *ex machina* as in Descartes. But Leibniz thought that Newton did, in fact, unwittingly fall victim to use of the *ex machina* device. Leibniz's own defense of something fundamentally active in nature was seen by him as a return to Plato and Aristotle, specifically to the ideas that physical existents were characterized by *kinesis* (motion) and *dynamis* (dynamic power). It even makes sense to claim that an account of the universe as dynamic and vibratory would have surprised Plato less than it would Newton (see Whitehead 1978: 94). And Kant appears to rebuke those who ridicule Leibniz's panpsychism because, *if* we had to state what physical reality is in itself, we would have to follow Leibniz (1900: Part 1, Ch. 1).

I am obviously not doing justice to the rich history of panpsychism in this brief introduction. But I am trying to militate against the famous (or infamous) quip allegedly made by Quine that there are two quite different reasons why people enter philosophy: to do philosophy or merely to do history of philosophy. I am arguing that these two are not mutually exclusive. That is, thinking about the mind-body problem *is* historical thinking via major figures like Plato and Descartes, whether philosophers of mind realize it or not. Regarding the mind-body problem in general, and regarding panpsychism in particular, the following well-known quotation from Alfred North Whitehead seems most appropriate: 'The safest general characterization of the European philosophical tradition is that it

consists of a series of footnotes to Plato' (Whitehead 1978: 39). In the present chapter I will be relying especially on the thought of Whitehead, Charles Hartshorne, and other process philosophers who have been quick to notice the panpsychist dimension of Plato's thought.

2. Some Panpsychist Passages

What does it mean to explain the world, in general, and human nature, in particular? On a Platonic basis such an explanation would seem to be in terms of either: (a) soul or (b) matter; or (c) both soul and matter. The second option would have been anathema to Plato. Further, there is considerable well-known evidence in Plato's dialogues in favor of the third option, leading many scholars to think of Plato as some sort of dualist. But the issue is complicated and not merely because of the understandable difficulty involved in determining Plato's own views in his dialogues. That is, there is also evidence in favor of the claim that the first, panpsychist option may be the best clue we have to understanding Plato's overall view. The main purpose of this chapter is to explore this option. In the course of this exploration I will also touch on the relationship between soul and Platonic forms as well as the importance of the World Soul in several of Plato's later dialogues. The forms and the World Soul make Plato's view somewhat unique in the history of panpsychism.

An initial passage to consider can be found at *Epinomis* 983d, where Plato's presumed spokesperson (the Athenian) suggests that soul is the universal cause of body. A key question is: how close do such passages take us to the claim that Plato was a panpsychist? His view seems to have been, at first glance at the *Phaedrus* (245e, also 246b, 275b) and *Laws* (896a, also 895c, 896d, 898d, 899a–b), that souls *initiate* change and transmit it to others, whereas bodies merely *receive* and transmit change. That is, soul is defined in terms of self-motion in the *Phaedrus* and *Laws* and, as a result, has a certain sort of ontological priority to body, a priority that is found in the *Epinomis* as well. (On the authenticity of the *Epinomis* as an epilogue to Plato's *Laws*, see Crombie 1962: 12, 329.)

For metaphysical or cosmological purposes, *psyche* is used by Plato to refer to experiencing, thinking, remembering, feeling, etc., and only for ethical or religious purposes does he use the term to refer to an enduring entity behind these processes. The dynamism of soul is evidenced in the new concrete reality that exists at each moment. Plato is well aware of the fact that the transition from motion to rest and from rest to motion does not occupy much of a stretch of time but occurs in an instant (*exaiphnes* – *Parmenides* 156). It is not illegitimate to infer that it is this self-moving, processual character of soul in Plato that is crucial in the effort to understand the relationship between soul and body.

An understandable mistake that could be made at this point would be to conclude that in Plato's dialogues soul is active whereas body is passive. But in the *Republic* it is clear that the degree of affective and cognitive adequacy in soul is due to the reliability of the influence it receives from the object that is experienced or known. This point enables us to see that it would indeed be a mistake to claim that in Plato souls are strictly active. In fact, one of the most obvious of the soul's powers is its ability to receive influence from others, as in perception. As one moves up the famous divided line in the *Republic* (509d–511e),

what one finds is increasing adequacy of apprehension of the world due to an increasing ability of soul to receive influence from the reality apprehended: from imagination (*eikasia*), to belief (*doxa*), to hypothetical knowledge (*dianoia*), to the highest level of awareness (*noesis*).

The point of the present chapter is not to claim that Plato has to be viewed as a panpsychist, but rather that one does not have to view him as a dualist, as is commonly assumed. A key passage to consider in the effort to make intelligible the panpsychist tendency in his thought is *Sophist* 247d–e (also see 249a, 249d), where Plato's presumed spokesperson (the Eleatic Stranger) says the following:

> I suggest that anything has real being that is so constituted as to possess any sort of power either to affect anything else or to be affected, in however small a degree, by the most insignificant agent, though it be only once. I am proposing as a mark to distinguish real things that they are nothing but power.

Whitehead goes so far as to say that in this passage can be found the height of Plato's genius as a metaphysician (Whitehead 1967: 120). Being is here defined as power or *dynamis*, which is not accidentally the root of our word 'dynamic,' specifically the power to affect, or the power to be affected by, others. I take it that the 'or' (*eite*) in the preceding quotation does not refer to mutual exclusivity between influence and being influenced, hence 'and' (*kai*) might have better expressed the point. Further, the words 'nothing but power' in the preceding quotation might be better phrased in terms of the capacity for *dynamis*.

Plato's dialogue style betrays this idea. One does something with (not to) one's dialectical partner. This makes the dialogue style a good model for the general nature of reality. As opposed to authoritarian dictation, in dialectic it will not suffice to suggest that to speak and be heard are admirable, whereas to listen and to hear are not. Plato makes this point explicit in at least one dialogue, the *Gorgias* (508a). We will also see that to speak of the supreme soul (i.e., the World Soul) as persuading other souls is to presuppose that each lesser soul has the power to be moved.

Plato hints at the panpsychist position (in which only concrete singulars feel, and in which the abstract is real only in the concrete, thus soul is the inclusive type of reality) when he indicates that soul is coincident with every action and passion. But no ancient Greek thinker was in a position to fully understand the difference between singulars and aggregates in the smaller parts of nature. That is, there is a vast difference between soul as such or soul as a generic principle, on the one hand, and animal soul (including human soul), on the other. It is the lack of self-motion in everyday inanimate things that has caused materialists or dualists to suppose that the microscopic parts of these things also lack self-motion (see Hartshorne 1983: ch. 2–3).

In Plato's dialogues we learn that soul is the universal cause (*aitias tou holou* – *Epinomis* 988d), that it is (metaphysically rather than chronologically) prior to body (*presbyteras e somatos* – *Laws* 892a), that bodies are derived from soul (*soma de deuteron te kai hysteron* – *Laws* 896c), that we receive our being from soul (*Laws* 959a), and that soul is the primary

source of all things (*psychen genesin hapanton einai proten* – *Laws* 899c). So although Plato could not fully understand the full significance of panpsychism in that he lived over two thousand years before the discovery of cells and other microscopic centers of power, it would be a mistake to think that he was totally ignorant of such significance by defending dualism *simpliciter*. (Also see *Philebus* 28d, 29a-31b; and *Timaeus* 30b–c, 31a, 40b–e, 69c–70e, 77b.)

William Lane Craig offers in formal outline something like the following argument (in abbreviated form), which indicates why the definition of soul in terms of self-motion is crucial in the effort to understand the panpsychist tendencies in Plato's thought. That is, a version of Craig's reconstruction of Plato's thought can be used not only to understand Plato's theology, but also to understand his flirtations with panpsychism:

a. Some things are in motion.
b. There are two kinds of motion: communicated motion and self-motion.
c. Communicated motion implies self-motion: (i) because things in motion imply a self-mover as their source of motion, otherwise there would be no starting point for the motion; and (ii) because things moved by another imply a prior mover.
d. If all things were at rest, only self-motion could arise directly from such a state: (i) because a thing moved by another implies the presence of another moving thing; (ii) but this contradicts the hypothesis that all things are at rest.
e. Therefore, the source of all motion is self-motion or soul

(Craig 1980: 4).

To say that there are strong grounds for thinking that Plato was a panpsychist is to say that self-motion is pervasive in Plato's cosmology, including the self-motion *of*, or self-motion *in*, plants, stars, the sun, and earth. That is, we would be wrong to assume that it is only human or animal souls that are capable of self-motion.

Although he finds the animism or panpsychism in Plato's dialogues (especially, the *Phaedrus*, the *Timaeus*, and the *Laws*/*Epinomis*) odd, Crombie emphasizes the importance of this aspect of Plato's thought. The self-motion of soul has not only intellectual, but also spiritual, significance in Plato's thought, as Crombie sees things, in that it is part of the effort to counteract the physicalist reductionism of the newly found atheists. The fact that Crombie himself does not find mechanism problematic is evidence in favor of the claim that he can be seen as an independent observer of Plato's animism or panpsychism (see Crombie 1962: 325–40).

3. The Forms and Soul

Two additional topics need to be considered in order to understand Plato's engagement with panpsychism. First, something has to be said about the relationship between soul and the forms in Plato. If the forms are seen as having causal power (and there is textual evidence that Plato viewed forms in this manner), then this might mean that whereas

bodies can be moved by other moved things or by self-moved souls, souls can also be moved by the motionless forms. Yet if the forms are items in divine *psyche*, then soul in some sense is fundamental for metaphysical speculation and hence Plato came quite close to panpsychism. That is, panpsychism is compromised if too much emphasis is placed on forms or abstract objects as having causal power of their own or as having agency.

It is common to hear that the forms are 'independent,' even of the divine soul. It must be admitted that there are several passages in Plato's dialogues that support this interpretation. But if 'X is independent of Y' has any sharp meaning it must be that X could exist even if Y did not, which implies that Y is contingent. If X stands for the forms and Y for divine soul, then the nonexistence of God is being taken as possible. But this result conflicts with the treatments of divine soul in both the *Timaeus* and Book Ten of the *Laws*. It also conflicts with Plato's anticipation of the ontological argument (see Dombrowski 2006). That is, if God's existence is not contingent in Plato, then not only are the forms envisaged by divine soul, they could not lack this status. Both the forms and God are everlasting on Plato's view and as a result there can be no real independence of one from the other. God always exists as the ideal knower in Plato and the forms are the objects such an ideal knower would know. In different terms, on the extradeical view (where the forms have some sort of curious existence apart from both matter and soul) panpsychism is threatened, but on the more defensible intradeical view (where forms are items in divine psychical process) panpsychism is preserved.

It can be said that the greatest problem in Plato's metaphysics does not concern the theory of forms, but rather concerns the problem of how to sufficiently grasp the functions of soul as both receptive and creative. There are also the related problems of understanding internal and external relations in soul and how to come to terms with the way that soul interacts with body. Plato's analysis of becoming remains incomplete, because if knowing something is to change that something, as Plato sometimes indicates, then past events go on changing when we think about them. Plato probably flirted with this idea (that knowing something changes it) as a reaction to the opposite view that the past completely determines the present, in souls as well as in bodies. The self-motion of soul must mean that soul originates change, which is at least compatible with the view that necessary, although not sufficient, causal conditions are inherited from the past. Soul does not merely transmit tendencies from the past, nor just receive them, as in bodies.

4. The World Soul

Second, it is difficult to understand the topic of panpsychism in Plato without a consideration of his cosmology, dominated as it is by the concept of a World Soul. The concept of soul as self-moving or self-creative sheds light on Platonic theodicy in the sense that the lack of complete order in the world is explained by there being many souls that can get in each other's way. These many self-active agents imply indefinitely great if not complete disorder unless there is a supreme Soul to persuade the many lesser souls to conform to a cosmic plan. But they cannot *completely* fit such a plan for then they would not be self-determined.

The meaning of God as the World Soul having power over us is intelligible only if we see the World Soul as a self-moved mover of others who is also partially moved by these other self-movers. Such a God can rule the world only in the sense that such a soul sets optimal limits for free action. The divine can control the changes in us by inspiring us with novel ideas, hence by molding the divine life itself God presents us at each moment with a partly new ideal. But such a God is not omnipotent because omnipotent power would be a monopoly of power over the powerless, which is at odds with the aforementioned Platonic view that being *is* power.

The doctrine of the World Soul appears in at least five of Plato's later dialogues (*Statesman*, *Philebus*, *Timaeus*, *Laws*, *Epinomis*), indicating its importance to Plato. To help explicate both the characteristics of the World Soul and the implications of the World Soul for panpsychism in Plato, three levels of *psyche* can be distinguished.

P1 is *psyche* at the microscopic level of cells and atomic particles, where contemporary biology and physics can be seen to have vindicated Plato's forays into panpsychism. Reductionistic materialism in its various manifestations and determinism have faded, on the contemporary panpsychist view (see Griffin 1998), as reality in its fundamental constituents itself seems to have at least a partially indeterminate character of (Platonic) self-motion. That is, the sum total of efficient causes from the past does not supply the sufficient cause to explain the behavior of the smallest units of becoming in the world. Plato was wiser than he knew. Little did he realize that in 20th–21st-century science universal mechanism might give way to ubiquitous self-motion.

P2 is *psyche* per se, which involves feeling of feeling, found in animals and human beings, where central nervous systems enable 'higher' organisms as wholes to feel just as the constituent parts show at least prefigurements of feeling at a local level. And feeling *is* localized. Think of a knife stuck in the gut of any vertebrate or of sexual pleasure. P2 consists in taking these local feelings and collecting them so that an individual as a whole can feel what happens to its parts, even if the individual partially transcends the parts. If the reader thinks that all of this is foreign to Plato, it should be noted that in the *Republic* (462c–d) Plato makes it clear that if there is pain in one's finger (note, not the whole hand), the entire community (*pasa he koinonia*) of bodily connections is hurt. The organized whole of the individual is such that when one part is hurt there is a feeling of pain in the human being as a whole (*hole*) who has the pain in the finger. Of course, Plato's treatment of localized pain was not based on cell theory, but on obvious phenomenological evidence symbolized in the *Republic* by precisely localized intense pain in a finger rather than in an arm or a hand. Panpsychism helps us to better understand how, when one has pain in one's finger, the pain is in one sense mine and in another sense not mine. Further, there was some sort of dim awareness among the ancient Greeks of nerves (*neura*), which were significant parts of the ancient scientific view of the world (see Solmsen 1961).

P3 is divine *psyche*. If I am not mistaken, the following four-term analogy takes us to the heart of Plato's view of both panpsychism and the World Soul:

P1 : P2 :: P2 : P3

The universe is a societal organism (a World Soul), of which one member (the Demiurge) is preeminent, just as human beings are societies of cells (*neura* or 'nerves'), of which the mental part is preeminent.

Because an individual must, to maintain integrity, adapt to the environment, mortality is implied. But if we imagine the World Soul we must not consider an environment external to divinity, but an internal one: the world body of the World Soul (including the demiurgic divine mind). This cosmic, divine animal (*zoon* – *Timaeus* 30c) has such an intimate relation to the divine body that there must also be ideal ways of perceiving and remembering the bodily parts such that the World Soul can identify the microindividuals (P2) that are included. We can only tell when cells in our finger have been burned by the fire. We cannot identify the microindividuals as such.

Although the evidence in Plato's dialogues is somewhat unclear as to how matter could consist of multitudinous souls of extremely subhuman kinds, he had at least a glimmering that it was the multiplicity of souls that made general order possible (due to the persuasive influence of the World Soul on lesser souls) and absolute order impossible (due to the self-moving power of each soul). Each new divine state in the life of the World Soul harmonizes itself with both its predecessor and with the previous state of the *cosmos*. This is analogous to a human being harmonizing itself with its previous experience and bodily state, but with a decisive difference. The human soul must hope that its internal and the external environment will continue to make it possible for it to survive, whereas the World Soul has no such problem in that there is no external environment for the World Soul. But the differences between the World Soul and the human soul (the World Soul knows the microindividuals included within the divine life and the World Soul has no external environment) should not cloud the important similarities (the fact that self-change is integral to soul at all levels and the fact that the soul-body analogy used to understand the World Soul-world body helps us to better grasp the relationship between a human soul and the besouled parts of a human being). The autokinesis of soul at all levels (P1, P2, and P3) is precisely what enables Platonic panpsychism to avoid the bifurcation of the world (and the reductionistic reaction to such bifurcation) that has dominated philosophy since the time of Descartes.

The most important similarity lies in the fact that one's bodily cells (or *neura*) are associated, at a given moment, with one as a conscious supercellular singular, just as all lesser beings are associated with the society of singulars which in the *Timaeus* is called the divine animal or the World Soul. In a way, all talk of the World Soul short of strict univocity contains *some* negativity, in that God does not exist, know, love, etc. exactly as we do. With regard to the divine body, however, almost all theists in the Abrahamic religions have allowed this negativity to run wild by completely denying the divine body. This is perhaps why the World Soul strikes many or most contemporary readers as odd.

It did not strike Plato as odd, however. In fact, according to Plutarch, all of the ancient philosophers, except for Aristotle and the atomists, believed that the world was infused with a divine, animal soul (Plutarch III 1870: 133). In two passages in Plato's later dialogues (*Theaetetus* 176b–c; *Timaeus* 90a–d; also see *Republic* 500c–d) it is suggested that the goal of human soul is to become as much like divine soul as possible (*homoiosis theoi kata to*

dynaton). This is possible because the World Soul and our souls are remotely akin, just as our souls are remotely akin to souls vastly inferior to our own, although these latter are at least proto-sentient centers of self-moving power. This *homoiosis* doctrine is at once metaphysical/cosmological, intellectual, moral, and aesthetic (from the Greek *aisthesis*, feeling). Plato's view strikes some people as odd (see Annas 1999; Mohr 1985) on the assumption that divinity necessarily involves otherworldliness or escape from this world. But belief in panpsychism and the related idea of a World Soul does not involve a flight beyond the natural world of becoming if the World Soul just *is* that world when considered as an integral, besouled whole.

5. Conclusion

It should now be clear why it is not necessary to view Plato as some sort of dualist. Scholars have always noticed the importance of *psyche* in the mesocosmos (P2). But they have paid insufficient attention to the macrocosmic significance of soul in Plato (P3). We are parts (P2) of the World Soul or of what is seen in the *Timaeus* as the divine animal (P3). This should give us a clue regarding the analogous part-whole relationship between the microcosmic parts (P1) that constitute us as mesocosmic wholes (P2). This clue has also largely been ignored by scholars, which perhaps explains why it is often assumed without argument that Plato was a dualist. The aforementioned four-term analogy actually runs both ways. Just as P3 helps to explain why there is a *uni*verse or a *cosmos*, in that the World Soul brings together what would otherwise be the scattered multiplicity of the bodily, so also there is a sort of *cosmos* or wholeness of an individual life (P2). That is, bodily life is *permeated by* besouled, dynamic power, hence it is a mistake to think of soul in Plato as having a sort of epiphenomenal existence hovering above body.

As was mentioned earlier, there is evidence in Plato's dialogues that is supportive of the claim that he can be viewed as a panpsychist. This evidence includes the ideas that: only soul is capable of self-motion; soul is the universal cause of body; soul is the only initiator of change in the universe; soul is both active and passive in the dynamic power it exhibits; and bodies are derived from soul, hence soul is metaphysically prior to body. The problem of how life (or consciousness) arises out of lifeless (or unconscious) matter is not one that even comes up on the basis of this evidence in that the dynamic power of being goes all the way down, on the view of Plato I am explicating in the present chapter.

On the panpsychist view, each event is a dynamic subject for itself and an object for others, hence no event is simply a subject or simply an object, contra dualism or materialism. That is, reality is characterized by the dynamic power in subjects that become objects. Interpreting Plato as a panpsychist thus enables us to better come to grips with a key feature of several of his later dialogues, which suggests that being has both an 'in itself' dimension and an 'in relation' character. *Psyche*, defined by Plato as self-motion, is required in order to understand the very dynamism of the world. This is because inactivity is inimical to soul, even when its amphibious qualities are taken into consideration: soul has one foot in the eidetic realm of changeless forms and one foot in the instantial camp of dynamic nature. Only soul has the right sort of contact to link up both the forms and

ubiquitous flux. Indeed, the Greek *physis*, which is usually translated into English as 'nature,' might more accurately be translated as 'process,' as Whitehead argues (1967: 150).

Granted, Plato is his own worst enemy in the sense that he did a very good job in providing evidence for the widespread thesis that he was a dualist. But three disparate positions are clearly defended in ancient philosophy by Plato and others which, when put together by later thinkers, can lead to a credible version of panpsychism: (1) Plato's discovery in the *Sophist* of the metaphysical concept that being *is* dynamic power to both exert influence on, and to be influenced by, others; (2) his definition of *psyche* in the *Phaedrus* and *Laws* in terms of self-motion, a definition that is amplified by Aristotelian *kinesis*; and (3) the Epicurean belief that ultimate reality is atomic in character, a belief that is amplified by the ancient Greek discovery of *neura*, as detailed by Solmsen (1961).

6. Postscript: Skrbina's Interpretation

It will be helpful to see the extent to which the view of Plato presented in this chapter agrees with the magisterial study of the history of panpsychism authored by David Skrbina (2005), who agrees that there is something problematic in the familiar conclusion that Plato was a dualist. There is also something problematic in seeing the forms as the ultimate realities in Plato, as we have seen previously in defense of the intradeical interpretation of the relationship between the forms and divine soul. It must be admitted, however, as Skrbina notes, that Plato can also be read as something of a reverse epiphenomenalist. Epiphenomenalism consists in the belief that body is the fundamental reality, whereas mind has some sort of shadowy reality that derives from the body. The reverse epiphenomenalism in Plato noticed by Skrbina, however, has mind as the fundamental reality, whereas the body has a shadowy sort of reverse doppelganger existence derived from mind, as found, say, in the famous myth of the cave in the *Republic* (see Skrbina 2005: 6, 10, 14).

Skrbina thinks that the best interpretation of Plato is that he was a panpsychist, especially when his later dialogues are considered. (Things get confusing, however, when Skrbina claims that the Stoics were the first panpsychists. I assume that this is a typographical error and that what Skrbina wanted to say was that the Stoics were the first *pantheists* –2005: 21.) Perhaps the biggest contribution that Skrbina makes to the view of Plato as a panpsychist is his argument that a belief in a panentheistic (not pantheistic) World Soul is different from a belief in panpsychism and that one could commit to a belief in a World Soul without committing to panpsychism in that to speak of soul/mind as a single universal being leaves open the question as to whether soul/mind is attributable to each thing in itself. Skrbina and I are in agreement that Plato can profitably be seen as *both* a defender of a panentheistic (literally, all is *in* God) World Soul *and* a panpsychist, given the proliferating pluralism of *psyche* in Plato (Skrbina 35, 39). However, Skrbina is for some unstated reason reluctant to view the World Soul in theistic terms.

Four arguments for panpsychism in Plato (and other ancient Greek philosophers) are detected by Skrbina. First, there is an argument from indwelling power in things (based on

the evidence provided by things being capable of self-motion and being possessive of other sorts of dynamic power). Second, there is an argument from continuity (apparently based on the idea that we can notice various levels of dynamic power in the universe and have a sense that such powers are not to be found solely in human beings). Third, there is an argument from first principles wherein soul/mind is not derivative or incidental, but central and primary. And fourth, there is the well-known argument from design that has exerted a huge influence on the history of philosophy, an argument wherein the self-motion of divine soul is transmitted to other besouled beings throughout the *cosmos*. These arguments, far from striking ancient philosophers as odd, would have had a family resemblance to the hylozoism or animism or panpsychism of the presocratic philosophers. In fact, the first philosophers who defended the idea of matter without feeling were the atomists, on Skrbina's interpretation. That is, what would have struck most ancient thinkers as odd would be the unfeeling atoms (*atomoi apatheis*) of Leucippus and Democritus (Skrbina 2005: 25–8, 33, 40).

From the preceding it would be legitimate to conclude that I see a great deal of compatibility between my own view of Plato's panpsychism and Skrbina's view, contra the interpretation of Plato as a dualist, as found, say, in Guthrie (1962–81: 4, 420). Or again, whereas Plato *did* tend toward dualism in his middle period, in his later dialogues he tended toward panpsychism, although I would add that there are prefigurements of panpsychism in the middle dialogues (as in *Ion* 533–536, where the magnetic power of lodestones is discussed) and faint traces of dualism in the later dialogues. But Skrbina is surely correct about the tendency toward panpsychism in the later dialogues, as in *Philebus*, where soul is the cause of all things and where there are living presences in nature (896d, 899b). This tendency in Plato is both noticed and affirmed by Plotinus (*Ennead* VI. 7, 11). Indeed, Skrbina insightfully goes so far as to say that 'There appear to be no passages in the late dialogues that explicitly deny the panpsychist conclusion' (2005: 44, also 34–5, 42–3).

Works Cited

Annas, Julia (1999). *Platonic Ethics, Old and New*. Ithaca: Cornell University Press.

Craig, William Lane (1980). *The Cosmological Argument from Plato to Leibniz*. New York: Barnes and Noble.

Crombie, I. M. (1962). *An Examination of Plato's Doctrines*, vol. 1. London: Routledge and Kegan Paul.

Dennett, Daniel (1991). *Consciousness Explained*. Boston: Little, Brown.

Dombrowski, Daniel (1991). 'Hartshorne and Plato'. In Lewis Hahn (ed.), *The Philosophy of Charles Hartshorne*. LaSalle, IL: Open Court.

Dombrowski, Daniel (2005). *A Platonic Philosophy of Religion: A Process Approach*. Albany: State University of New York Press.

Dombrowski, Daniel (2006). *Rethinking the Ontological Argument: A Neoclassical Theistic Response*. New York: Cambridge University Press.

Griffin, David Ray (1998). *Unsnarling the World Knot: Consciousness, Freedom, and the*

Mind-Body Problem. Berkeley: University of California Press.
Guthrie, W. (1962–81). *History of Greek Philosophy*, 6 vols. Cambridge: Cambridge University Press.
Hartshorne, Charles (1983). *Insights and Oversights of Great Thinkers: An Evaluation of Western Philosophy*. Albany: State University of New York Press.
Kant, Immanuel (1900). *Dreams of a Spirit-Seer*. Trans. Emmanuel Goerwitz and Frank Sewell. New York: Macmillan.
Kim, Jaegwon (1993). *Supervenience and Mind*. New York: Cambridge University Press.
Leibniz, Gottfried (1991). *Discourse on Metaphysics and Other Essays*. Trans. Daniel Garber and Roger Ariew. Indianapolis: Hackett.
McGinn, Colin (1991). *The Problem of Consciousness*. Oxford: Blackwell.
Mohr, Richard (1985). *The Platonic Cosmology*. Leiden: Brill.
Nagel, Thomas (1986). *The View from Nowhere*. New York: Oxford University Press.
Plato (1977). *Platonis Opera*. Ed. John Burnet, 5 vols. Oxford: Clarendon Press.
Plato (1999). *The Collected Dialogues of Plato*. Ed. Edith Hamilton and Huntington Cairns. Princeton: Princeton University Press.
Plutarch (1870). *Plutarch's Morals*. Ed. William Goodwin. Boston: Little, Brown.
Royce, Josiah (1908). *The Philosophy of Loyalty*. New York: Macmillan.
Seager, William (1991). *Metaphysics of Consciousness*. London: Routledge.
Seager, William (1999). *Theories of Consciousness*. London: Routledge.
Searle, John (1992). *The Rediscovery of the Mind*. Cambridge: MIT Press.
Skrbina, David (2005). *Panpsychism in the West*. Cambridge: MIT Press.
Solmsen, Friedrich (1961). 'Greek Philosophy and the Discovery of Nerves'. *Museum Helveticum*, 18: 150–67, 169–97.
Strawson, Galen (1994). *Mental Reality*. Cambridge: MIT Press.
Strawson, Galen (2006). *Consciousness and Its Place in Nature: Does Physicalism Entail Panpsychism*? Exeter: Imprint Academic.
Whitehead, Alfred North (1967). *Adventures of Ideas*. New York: Free Press.
Whitehead, Alfred North (1978). *Process and Reality*. Ed. David Ray Griffin and Donald Sherburne, corrected edition. New York: Free Press.

3
ABHIDHARMA PANPROTOPSYCHIST METAPHYSICS OF CONSCIOUSNESS

Monima Chadha

Panpsychism refers to a variety of doctrine that asserts that mental features are ontologically fundamental and ubiquitous. This definition in terms of mental features is not standard, but I think it will prove to be useful. The term 'mental features' is suitably ambiguous: it is neutral on whether mental features are restricted to phenomenal features or comprise non-phenomenal features as well and also on whether mental features are substances, or objects, or properties or some entirely new kind of thing hypothesised by physicists or contemplated by metaphysicians. To claim that mental features are ontologically fundamental separates out panpsychism not only from physicalism but also from eliminativism. Again, to claim that they are ubiquitous separates out the doctrine from Cartesian dualism and neutral monism. However, this fails to rule out idealism as a variety of panpsychist views. I think rightly so, as idealism is a version of panpsychism, indeed its most popular version in the nineteenth century (Seager and Allen-Hermanson 2015). The revival of panpsychism in recent times is not a version of idealist metaphysics; far from it. Most contemporary panpscyhists accept that the material and mental features are ontologically fundamental. Some, like Galen Strawson (2006), go so far as to claim that panpsychism is entailed by 'real physicalism' or 'realistic materialism'. Much more needs to be said to give a precise specification of panpsychism but this brief characterisation will suffice for my purposes here. I do not intend to give a precise answer to the question as to what is panpsychism; rather my interest here is to explore what Buddhist views of the mind have to offer to contemporary debates on panpsychism.

Contemporary interest in panpsychism is driven by the intractablilty of the "hard problem" (Chalmers 1996). Reductive physicalism is no longer the only game in town, even scientifically minded philosophers are willing to pursue other paradigms as live options. As a methodological approach, reductive physicalism, or the so-called orthodox naturalism of "eliminating or locating" (Jackson 1998: 5) non-physical features of the world in the physical, has given way to liberal naturalism (Strawson 1985). Liberal naturalists in

philosophy of mind seek to explain consciousness without the metaphysical constraints of physicalism and of any ontological limitations on sorts of things that constitute the basic furniture of the world. Panpsychism as a prominent example of the liberal naturalist approach has the license to introduce a different kind of metaphysics to replace the current physicalist ontology and its settled categories. Panpsychists should not be restricted by ontological categories like those of substances, properties, atoms, subjects of experience and so on. Contemporary panpsychists, for example Nagel (1998) and Strawson (1999), contend that our current conceptions of mind and body, mental and physical are radically inadequate and in need of revision. In light of this, it could be useful to look at philosophical traditions other than our own Western tradition, particularly at traditions that have grappled, and continue to grapple, with cognate problems. I believe that the ancient Indian Buddhist tradition, and its contemporary philosophical forms, is quite germane here for several reasons.

First, Buddhist philosophy contains a wealth of material that is relevant to the concept of consciousness and that of mind. Panpsychists, and analytic philosophers more generally, can find valuable insights in it. However, Buddhist philosophy, and classical Indian philosophy more generally, is not easily integrated into mainstream Western philosophy because their interests (specifically, soteriological aims) and languages are foreign to contemporary philosophers and have been ignored for a very long time. It is only in the last decade or so that scholars, with the requisite philosophical training and knowledge of the relevant languages, have successfully extracted discussions of consciousness and mind from classical Buddhist sources and situated them in the corresponding contemporary philosophical debates (Dreyfus 2011; Thompson 2014; Ganeri 2012; Chadha 2015; Garfield 2015). Secondly, what makes the Buddhist tradition especially attractive to panpsychism as well as liberal naturalism is because its metaphysics is radical and revisionary in that it lacks a self and any other persisting entities. Thirdly, and most importantly, the reason in its favour is that panpsychism seems to have a natural affinity with Buddhism more generally as both views emphasise that the mental is *almost* a universal feature of living beings. This statement is possibly true but only because it is vague and imprecise. Buddhist philosophers will complain that Buddhist metaphysics contains a variety of views about the nature of mind, some of which clearly are not panpsychist. Garfield (2015), for example, argues that Buddhist texts can be of great use to contemporary physicalism since they contain insights that can be used to respond to the famous conceivability argument, in favour of dualism, for the possibility of zombies. Panpsychists will complain that 'almost' is not a part of the literal meaning of panpsychism. My response is that my interest is in elucidating *a* panpsychist Buddhist view of consciousness to explore ways of strengthening and enriching contemporary panpsychist positions in important ways by drawing on Buddhist thought. My response to the Buddhist is that there is no *one* view that qualifies as *the* Buddhist view of consciousness. Therefore, what I shall be exploring here is *a* Buddhist view of consciousness and this should answer any concerns of Buddhist philosophers. My focus will be on the Abhidharma Buddhist tradition, which flourished in the first millennium AD in India, and whose stated aim is to give an account of conscious experience. In particular, I want to draw attention to the view of consciousness defended by Sautrāntika-*Yogācāra*

School, or what is commonly called the Buddhist Logic School (Siderits 2007: 208). This view is originally proposed by Vasubandhu in his formulation of the Sautrāntika doctrine in the *Abhidharmakośa-bhāsya*, and later modified by rich insights from *Yogācāra* philosophers like Asanga, Dignāga and Dharmakīrti. The challenge for them is to give an account of conscious experience in the absence of self and indeed all other persisting and substantial entities. I will show that the Sautrāntika-Yogācāra view of consciousness presents a plausible version of panpsychism. And, that it can address some longstanding concerns about panpsychist theories, especially the combination problem. Furthermore, recent neuroscientific studies of Buddhist meditation practices give us some independent, some would say empirical, reasons to think that these views of the mind are plausible.

The plan of this chapter is as follows. In section 1, I present the Sautrāntika-Yogācāra view of consciousness. In section 2, I argue that this view, as a version of panprotopsychism, has the potential to offer an adequate solution to a major challenge that continues to plague various panpyschist views: the combination problem. To be clear, my aim is not to offer an Abhidharma Buddhist solution to the combination problem, but only to show how a panprotopsychist view might explore solutions to the combination problem. In his papers (Chalmers 2015, 2016) Chalmers lists various versions of panpsychism and the combination problem. These two papers together chalk out a conceptual map carving out the logical space into a set of locations, positions, or places occupied by various panpsychist views and the solutions to the combination problems explored by their advocates. One of the tasks of this chapter is to place the Abhidharma Buddhist view on this conceptual map as a first step in exploring an Abhidharma Buddhist-inspired solution to the combination problem. This is interesting because it offers an insight into what protophenomenal properties look like and how we may try to overcome the non-phenomenal/phenomenal gap which would be the most serious combination problem threatening panprotopsychist views.

1. Sautrāntika-Yogācāra Model of Consciousness

The Sautrāntika-Yogācāra is part of the Indian Abhidharma tradition in Buddhist philosophy. Like all other Buddhist schools, the Abhidharma schools endorse the no-self doctrine. In fact, the most important Sautrāntika-Yogācāra philosopher Vasubandhu endorses a really strong version of the no-self view: the no-subject view (Chadha 2015). The challenge for the Abhidharma philosophers to explain the phenomenology of experience in a selfless, or more precisely for Sautrāntika-Yogācāra in a subjectless world.

The Abhidharma analysis of experience reveals that what we experience as a temporally extended, uninterrupted, flow of phenomena is, in fact, a rapidly occurring sequence of causally connected events each with its particular discrete object; much the same way a rapidly projected sequence of juxtaposed discrete images is perceived as a movie. Sentient experience is explained in terms of mental and physical processes (*skandhas*) that arise and cease in a causal sequence.[1] The distinctive contribution of the Ābhidharmikas is that they reduce the time scale of sequential mental and physical processes and regard them as discrete momentary events or *dharmas* (Ronkin 2005: 66–78). These discrete momentary events (*dharmas*) exist only for an instant or moment (*ks˙an˙a*), where a moment is the

limit of time. There are various opinions regarding the duration of a moment. Vasubandhu says that there are sixty-four moments in the time that it takes for a healthy man to snap his fingers (*Abhidharmakośabhāyam* 3.85b-c); others even say that billions of mind moments elapse in the time it takes for lightning to flash or eye to blink. These discrete momentary events or '*dharmas*' are the only existents, and they alone are objects of knowledge. Abhidharma produces a complete account of reality in terms of this single category of ultimately real entities. In this Section, I first explain the Sautrāntika doctrine of nature of *dharmas* and then explain how they combine to constitute conscious experiences.

The Sautrāntika view of the nature of *dharmas* and the constitution of conscious experiences from the basic *dharmas* arises as a result of a series of critiques of the early Abhidharma Sarvāstivāda doctrine. Vasubandhu rejects the distinction between substance and attributes (Pruden 1988: 1348). It is wrong to think of *dharmas* as substances, since substances are independent existents. However, according to the Buddhist doctrine of dependent origination (*Pratītyasamutpāda*) there are no independent existents; everything that exists is conditioned by its causes. Furthermore, *dharmas* are not universals, for Vasubandhu argues against the existence of universals (2.41a). Also, *dharmas* are not composite entities: they are simple, without spatial or temporal parts. Vasubandhu argues that medium-sized physical objects that we normally regard as real like jars, heaps, streams, persons, etc., are composites or aggregates of *dharmas*. These composites, he argues, do not exist in the ultimate sense, they are real only with reference to conception (*prajñapti-sat*); only *dharmas* exist in the ultimate sense (*paramartha*-sat) (6.4). But, then, one may ask: what are these *dharmas*, these ultimate constituents of reality? The Sautrāntikas have very little to say on this. They claim that *dharmas* are of two kinds: conditioned and unconditioned.[2] Among the conditioned *dharmas*, there are material (*rūpa*) and mental (*citta*, *caitta, etc.*) *dharmas*. Material *dharmas* (*rūpa*) are in space, but consciousness 'not having a mass, is not situated in a place' (1.43a–b).

Contemporary Buddhist philosophers have proposed that the best way to understand a *dharma* is to think of it as 'an elementary quality or event or condition' (Goodman 2004: 393) or 'a thin property' (Ganeri 2001: 99); a 'trope' as in contemporary analytic philosophy: e.g., the particular white of this wall rather than whiteness as such. Tropes, like physical objects, are particulars, as they are located in a particular time and space, though not necessarily so. They can be individuated as primitive particulars without spatio-temporal dimensions (Maurin 2014). They are unlike universals, in that they are not wholly present in each instance although they point to groups of identical things. However, in some sense, they share features with universals, in that they are 'ways objects are'.[3] Trope theorists believe that both particulars and universals can be understood as aggregates or sets of tropes. So, for example, we can think of Vasubandhu as an aggregate of Vasubandhu's colour, Vasubandhu's shape, Vasubandhu's logical acumen, etc. and similarly we can think of intelligence as bundle of Vasubandhu's intelligence, Asanga's intelligence, Udayana's intelligence, etc. Likewise, the Abhidharma would analyse the middle-sized mental and physical objects as aggregates of *dharmas*. Thus, we can think of mental *dharmas* as tropes that have temporal dimensions only.[4]

Such a trope-theoretic interpretation of *dharmas* has been used to reconstruct an Abhidharma account of the physical world (Goodman 2004) as well as of the conscious experience (Ganeri 2012) by an appeal to the notion of Humean supervenience (Lewis 1986: ix). Humean supervenience is the view that everything supervenes on spatio-temporal distribution of intrinsic properties. Weatherson (2015) offers a useful picture to understand this thesis. Imagine that the world is like a giant video monitor. The facts about a monitor's appearance supervene, plausibly, on intrinsic qualities of the pixels, together with facts about the spatial arrangement of the pixels. However, I want to steer away from Humean supervenience in offering a reconstruction of conscious experiences for two reasons: (a) *dharmas* are not necessarily physical, and for this reason (b) they are also not necessarily spatial entities (see Pruden 1.43 a–b previous). Thus, Humean supervenience, which necessarily involves intrinsic properties instantiated at spatio-temporal points, is not suited to an Abhidharma analysis. My own reconstruction will use the notion of *dharma*-clusters, vertical causation and supervenience.

There is intense debate about whether the Ābhidharmika philosophers arrived at this view through logical analysis or through introspection. Ganeri favours the former in suggesting that it is best to think of *dharmas* as proto-intentional (proto-cognitive and proto-affective) psychologically primitive processes that combine to constitute conscious experiences (2012: 130). Dreyfus seems to favour the latter though with an important qualification. He notes that "the Abhidharma analysis of the mental is a description of the complexity of the components of mental processes as they are phenomenologically available, not an analysis of the ontological basis of mental processes, a basis that is not readily available to the kind of analysis central to its project" (Dreyfus 2011: 119). *Dharmas* can be thus thought of as the basic components of mental and physical processes, basic in the sense that they cannot be further analysed into more basic phenomenologically available components. Dreyfus adds the qualification that this should not be taken to mean that *dharmas* are directly given in introspection, for they may only be available to those who have had appropriate mental training, for example in mindfulness meditation. Note, however, that these two views are not mutually exclusive: we can think of *dharmas* as basic components that can be arrived at either through logical analysis or through mindfulness meditation training or through both; in fact Vasubandhu, in his magnum opus *Abhidharmakośa*, endorses precisely this dual route.

For Buddhist epistemology generally, sensory perception is the paradigm of conscious experience. According to most Abhidharma schools, sensory perception is always intentional, and is brought about by an interaction among the sense faculties (e.g., eye), the corresponding type of consciousness (e.g., visual consciousness) and their appropriate sense objects (e.g., form, colour, etc.). There are six kinds of consciousness: five corresponding to the sensory organs and the sixth the mental awareness (awareness of thoughts, feelings, etc.). Conscious experience, according to the Ābhidharmikas, is always directed at objects, i.e., it is *of* something. In this the Abhidharma philosophers agree with phenomenologists that intentionality is an essential feature of consciousness. What distinguishes the Abhidharma view is its analysis of object-directedness involving five universal mental features that accompany every conscious mental state. The role of these mental features in

conscious experience becomes obvious if we take into account the fact that the mere coming together of an object and sense faculty is not sufficient for a conscious experience to arise. A cognitive event is initiated by contact, which the Ābhidharmikas describe as a relation between a sense faculty and sense object giving rise to a sensory consciousness. Such sensory consciousness, according to these philosophers, is always associated with affect or feeling. The object may be felt as pleasant or unpleasant or neutral, but it is never sensed without arousing feeling. Perception is the mental factor which plays the role of discerning or discriminating the object by distinguishing it from other things. Furthermore, each conscious state is also goal-directed in that it is always associated with volition or an intention to act. This together with feeling is responsible for determining the ethical quality of consciousness, that is to say, whether it decreases or increases suffering. Attention is the mental factor that is responsible for orienting consciousness towards its object, in that it guides other mental features towards the object of consciousness (Thompson 2014: 38). The senses always process a steady stream of sensory impressions; some clusters of these impressions generate a representational form of the object when attention directs the universal mental features towards it. Attention thus is necessary but not sufficient for conscious experience. Representational forms of experience are created by the interplay of these five mental features and occasionally involve some other associated mental factors. The Abhidharma has a long list of these occasional mental factors; indeed, there are significant disputes on exactly how many there are. A few of these occasional features are: anger, greed, mindfulness, compassion, wisdom, etc. Each of these features when accompanying a conscious state has the ability to affect the way we see objects and thus contribute to its representational form. Ganeri explains this succinctly:

> The great elegance and attraction of the [Abhidharma] theory lies in the fact that simultaneously it recognises the irreducibility of the phenomenal character of experience, it admits the joint contribution of sensation and conceptualisation in the constitution of experience, it acknowledges that experience is, as it were, saturated with affect, that appraisal is built into the fabric of experience, it maintains that every experience has, as a basic ingredient, a capacity or tendency to combine in various ways with various others, and it makes the attention intrinsic to experience.
>
> (2012: 127)

The Yogācāra account of mind adds two further consciousnesses to the six kinds of conscious states accepted by the Abhidharma schools mentioned earlier. These are the basic or storehouse consciousness (*ālaya-vijñāna*) and the afflictive mentation or ego-consciousness (*kliṣṭa-manas*). The first is a constant and neutral baseline consciousness that serves as repository of all basic habits, tendencies, and karmic latencies accumulated by the individual. The doctrine of basic consciousness is in large part an attempt to show that there is mental continuity within a lifetime and across lifetimes despite the *dharma* ontology in which there is no enduring substance.[5] The ego-consciousness, on the other hand, is an innate sense of self arising from the apprehension of basic consciousness as being a self (Dreyfus and Thompson 2007: 97). This self, however, is not an ontological reality: it is merely a conceptual fabrication resulting from the (mis)apprehension of basic consciousness.

In addition to these two new kinds of consciousnesses the Abhidharma-Yogācāra *philosopher* Dignāga *introduces the notion of 'self-awareness' in Pramān˙asamuccaya*, his *seminal work*. He defines self-awareness as a mode of awareness that provides immediate, non-conceptual access to how things subjectively appear to the mind. Dignāga was the first to articulate the idea that consciousness requires reflexive awareness. In Buddhist philosophy of mind the term "consciousness" does not stand for a faculty or a mysterious property of physical matter or non-physical substance. Consciousness is simply a conscious state, which for the Buddhist-Abhidharma philosopher is composed of a complex network of conscious momentary atoms (*dharmas*). Reflexive awareness can be understood as a kind of mental perception. It is "mental" in the sense that it is independent of the five external senses and it is "perception" in that it is an immediate and therefore non-conceptual awareness. And, insofar as it is a perception, self-awareness is a means of valid cognition (*pramān˙a*). Dharmakīrti explains this idea by appeal to a representational theory of perception: perception of an object is mediated by the direct apprehension of the mental representation that reveals an aspect (*ākāra*) of the object (Dreyfus 1997). Although the form of an object is produced by a causal link between the object and a sense organ, the form itself is a feature of the experience, not a feature of the object of experience. This is the unique Buddhist view that the representational form of the object is also the phenomenal form of the experience. That is to say, the mediating representation is Janus-faced. Dignāga's dual-aspect theory can be brought in to explicate this further. On this view perception is a cognitive episode that grasps an object, e.g., a glass of wine. This perception in its arising is associated with a pleasant feeling which is immediately grasped as what-it-is-like for me to perceive the glass of wine. Thus, this single cognitive episode must involve two aspects: the objective form and the subjective form. The first that is the objective form (*grāhyākāra*) is the representation of the external object and the second that is the subjective form (*grāhakākāra*) is the grasping or holding of this representational or objective form. The phenomenal form that reveals itself in self-awareness. Dharmakīrti expresses this by saying that when we are aware of something we are simultaneously aware of the awareness (*PV*, III: 266).[6] Basic consciousness and reflexive self-awareness are essential to explain full-bodied conscious experience, whereas ego-consciousness plays the role of explaining away the imaginary self. Let me use an example to lay out the details of the Yogācāra-inspired account of conscious experiences.

Consider a present conscious experience of having a coffee. The *dharmas* that constitute aroma, colour, warmth, etc., are simultaneously provided by various sense consciousnesses. But there are countless many other coffee-relevant *dharmas* available in the basic consciousness that have resulted from previous encounters with coffee, for example, *dharmas* of taste, the caffeine rush associated with drinking coffee, the smell of freshly roasted coffee beans, the sweetness of added sugar, the desire for coffee, etc. The conscious experience of coffee does not arise in the first moment of sensory contact with the coffee; rather, this sensory contact sets off a chain of horizontal and vertical processes that might, depending on the availability of universal and occasional mental factors, result in the production of the representational form of the coffee. The presentation of coffee gives rise to sensible qualities that constitute aroma, color, warmth, *and* so on. These sensible

qualities horizontally cause contiguous corresponding (aroma, color, warmth, etc.) *dharmas* in the next moment. At the same time they vertically cause processes in the basic consciousness, which activate coffee-relevant *dharmas* (for example, desire to drink coffee, taste, the accompanying rush of caffeine, sweetness of sugar, etc.). These vertical activations produce new horizontal series. At a unique and propitious moment when all the relevant *dharmas* are co-present together with the requisite universal mental features (and other occasional mental factors) the representational form of the coffee is produced. Intention, although itself a universal mental feature and thus a *dharma* has a double role: it binds together the relevant *dharmas* to produce the representational form the object.[7] Attention too in the same way has a double role: it transforms this representational form into a phenomenal form. For these Abhidharma philosophers, conscious attention is top-down and conscious phenomenon, though intentional, may be partly unconscious. Conscious attention is what explains the for-me-ness of an experience. In other words, the production of the phenomenal form, which encompasses the subjective and the objective aspects, coincides with the awareness of the coffee and the awareness of the perception of the coffee. To summarize: a conscious experience of coffee and the awareness of that experience are simultaneously produced as a result of the production of the phenomenal form of the object, which supervenes on co-temporal *dharmas* that are vertically present at that moment.

There is evidence to suggest that such a picture of conscious experience is empirically plausible. Our sensory system constantly receives multiple inputs, which are usually perceived as a stream. Thus, perception is regarded as a continuous process. However, there is scientific evidence to believe that perception consists of discrete mind moments. Recently Baumgarten et al. (2015) have shown that neuronal oscillation cycles define discrete perceptual moments which constitute the basis for a discontinuous and periodic nature of somatosensory perception exactly as the Abhidharma believed. Furthermore, it makes sense if a brain/mind embeds within it what Varela et al. (2001) called a brainweb, a massive parallel distributed system of highly specialized processors. The "global neuronal workspace model" for conscious access imposes a temporal granularity on neural states in a stream of consciousness. This framework postulates that, at any given time, many modular cerebral networks are active in parallel and process information in an unconscious manner. Information is consciously represented when the neural population that represents it is mobilized by top-down attentional amplification into a brain-scale state of coherent activity (in Buddhist terms, the transformation of representational form into the phenomenal form) that involves neurons distributed throughout the brain. Global Workspace Theory (Baars 1988; Dehaene and Naccache 2001) suggests that consciousness does not make more information available; it just makes the representation available to multiple networks to cooperate and compete in solving problems. There is no central processor or subject that is conscious of the information. The representation itself is conscious merely because of its salience in the cerebral networks. This completes my account of Sautrāntika-*Yogācāra* view of conscious experiences.

2. What Kind of Panpsychist View Is Offered by Sautrāntika-Yogācāra?

The foregoing section was meant to give the reader a sense of the panpsychist flavour of the Sautrāntika-*Yogācāra* analysis of conscious experience. The fact that mental *dharmas* are part of the basic furniture of the world, along with physical *dharmas*, naturally leads to the conclusion that the Sautrāntika-*Yogācāra view is a version of panpsychism or panprotopsychism* (Chalmers 2015). If we think that conscious states supervene on collections of present mental *dharmas* which are best thought of as proto-conscious or proto-intentional features as the result of logical analysis, then we favor the panprotopsychist option. Alternatively, if we hold that mental *dharmas* are potentially phenomenologically available as they can be discerned as such by experts who have mastered the art of mindfulness meditation, then we favour the panpsychist option. For my purposes here, I shall concentrate on the panprotopsychist option as it is more amenable to analytic philosophy. The panpsychist interpretations requires the inculcation and perfection of meditation skills which is too much to expect from ordinary folk and philosophers.

The Abhidharma phenomenology is unique in that it presents an analytic method to discern, list, and classify the momentary mental *dharmas*. It advances a model of mind, of mental states – conscious experiences, thoughts, etc. – in delineating the components of mentality and how they interact to produce conscious experiences. However, it is not a fixed list. Different Abhidharma thinkers and traditions offer different lists and leave open the option that the list may be expanded and reclassified by later thinkers. Chalmers (2015) notes that we must wait for a full panprotopsychist theory to get clear about the nature of protophenomenal properties. As a version of the panprotopsychist view, the Sautrāntika-Yogācāra gives us some idea of what protophenomenal properties could be like. More generally, Abhidharma philosophies offer detailed lists of the sort of mental features that constitute conscious experiences.

Mental *dharmas* are best understood as entities that play a constitutive role in the production of conscious states. *They are classified into types according to the role they play in the constitution of conscious experiences and thoughts.* The Abhidharma does not provide a fixed and final list of types of *dharmas* as a result of the reductive analysis. Rather each of the traditions offers a list of irreducible elements but these lists are open ended and not immune to revision. The Theravāda school of Abhidharma introduced a system of eighty-two *dharma* types, the Sarvāstivāda school, on the other hand, adopted a system of seventy-five basic types of *dharmas*. The teaching of Abhidharma is a kind of therapy for those who are confused about the nature of reality. To begin with the Abhidharma teaching focuses on ridding oneself of the false sense of self that is imposed on what really exists. Once we discharge the enduring sense of selfhood in favor of a dynamic system of constantly changing and interrelated dharmas, we undermine the fruitless activity of trying to grasp and fix the world of experience. The reexamination of basic features that constitute the world of experience through proliferating lists is itself a method for challenging our yearning for a fixed and stable sense of the world (Heim 2014). On this reading the Abhidharma view tells us that the variety of experience is explained by a relatively small number of *dharma* types. This is in agreement with what physics tells us about the basic ingredients of the world. Not only this, the fact that the Abhidharma attitude is that these

lists are open-ended and revisable and that there is no "fixed" number of basic ingredients that describe the experiential world once and for all is very much in agreement with what we may call the scientific attitude.

Furthermore some of these mental features are universal in that they accompany every mental state; others are occasional mental features that accompany some thoughts. The occasional features are further classified into those that accompany "good" mental states, those that accompany "bad" mental states, and yet others that are indeterminate in that they accompany "neutral" states. The qualifications "good", "bad", etc., may seem odd to contemporary Western philosophers, but it must be remembered that in the Buddhist way of thinking most conscious experiences and thoughts are said to contribute to, and thus be responsible for, moral agency and actions. The moral quality of the mental state is determined by the quality of the constituting *dharmas*: "Good" mental *dharmas*, e.g. the feeling of compassion, the intention of sharing, etc., result in good mental states. "Bad" mental *dharmas*, greed, anger, etc., on the other hand, result in bad mental states. The Buddhists have a very simple way of measuring good, bad and neutral. That which reduces suffering is good, that which increases suffering is bad, and that which does not alter the suffering is neutral. Apart from these morally loaded *dharmas*, the Abhidharma lists also include: sensuous properties like redness, roughness and mellowness, etc.; feelings like joy, pride, etc.; intentions like good, bad, etc.; kinds of attention, meta-attention, calmness, lightness of body; and, so on. This very brief description is by no means a complete census of the many mental features that can be involved in conscious awareness. The task here is to give the reader a sense in which protophenomenal properties contribute to, and play a role in, the constitution of phenomenal properties. Protophenomenal properties determine the nature and character of phenomenal properties; they are not merely structural properties.

At this point, I want to turn my attention to the nature of combination among the *dharmas*. It is plausible to say that *dharmas* can combine with one another, but a lot turns on whether the rules for this combination are laws of nature (*a posteriori*, metaphysically contingent) or inherent in the nature of the combining things (*a priori*, metaphysically necessary). Panprotopsychism has an advantage over emergentism provided that the rules of combination are of the latter type (*a priori* and necessary), otherwise the view cannot be differentiated from emergentism. We have already mentioned that the Abhidharma ontology includes the following universal mental factors which must accompany all conscious experience: contact (*sparśa*), feeling (*vedanā*), perception (*sam ̇jñā*), intention (*cetanā*) and attention (*manasikāra*). The last two – intention and attention – are special factors. Intention, although itself a *dharma*, is characterised in the Theravāda as something that through its activity brings together the other mental features to produce the objects of experience as representational forms (Heim 2014: 105). Representational forms as objects of experience do not mirror the world out there; rather they are actively constructed by the cooperation of mental factors and external objects. Attention is what leads the mind to its object. An analogy used by Buddhaghosa suggests that attention is like a coachman, it drives other mental factors to an object (*ibid.*). Attention is what gives rise to the phenomenal form of the object or the subjective aspect of the experience. Reflexive

awareness then provides the principle that ties together the subjective aspects of experience to give rise to the self-conscious states (Ganeri 2012). The important point to note here is that intention, attention and self-awareness are essential and intrinsic features of experience itself. Thus according to the Buddhist panprotopsychist, Abhidharma theories of the rules for combination are inherent in the nature of the combining things, thus they are metaphysically necessary and *a priori*. This ensures that Buddhist panprotopsychism is not vulnerable to the sort of objections that may be raised against the emergentists.

Panprotopsychism has some advantages over its rivals. Chalmers, for example, notes that in contrast to other panpsychist views, panprotopsychism does not need to posit subjects or proto-subjects at the bottom level. But it does suffer from the especially difficult combination problem that he labels the "non-phenomenal/phenomenal gap." This version of the combination problem, like the corresponding problems for other versions of panpsychism, is spelt out by means of a conceivability argument. The weight of the argument rests on the existence of protophenomenal zombies: beings that share our protophenomenal properties (and also all the physical properties) but can still be without phenomenal consciousness. Chalmers admits that the conceivability of protopanpsychist zombies is less obvious than the possibility of panpsychist zombies (2015). The foregoing Buddhist account gives us some idea about the protophenomenal elements that constitute phenomenal states of conscious awareness. This random list might lead one to question whether panprotopsychist zombies are really conceivable, a suspicion strengthened by the Abhidharma explanation of subjectivity without a self. As noted earlier, the Buddhist account does not accept that the usually unargued assumption that having a conscious experience necessarily entails a subject of experience, a subject for whom it is, somehow or other, like to have this experience. On this Buddhist account subjectivity is accounted for by the fact that conscious states are reflexively self-aware, they are not states of awareness in virtue of being owned or had by a subject of experience. Reflexivity is the very nature of consciousness; it defines what it is to be conscious. The no-self thesis is a corollary not only of the denial of persisting entities, but also of the denial of a substantialist metaphysics of independently existing entities. This very minimal metaphysics of momentary *dharmas*, property-particulars, eliminates the subject of experience completely. There are no subjects and no streams of experiences; only synchronically unified experiences. Such a denial of subjects of experience at the level of microexperiences and macroexperiences is much more radical than contemplated by Western philosophers, with the possible exception of Hume. This view of continually evolving causal series of momentary tropes has no room for the so-called non-subject/subject gap that is offered as a basis to argue for the plausibility of existence of panprotopsychist zombies.

Someone may object[8] that the Abhidharma are not really defending a no-subject view, rather they are only denying an eternal, metaphysically fundamental, incorruptible soul substance. Those who are concerned about the combination problem do not have such a subject in mind. The objection is as follows: let the subject be whatever it is that accounts for the subjectivity of conscious states. If reflexive awareness is what accounts for the subjectivity of conscious states and reflexive awareness is itself a *dharma*, then we do need to combine subjects after all. I think the Abhidharma must at this juncture claim that not all

mental *dharmas* are reflexively aware and thus should be treated as subjects. The point is that reflexive awareness is triggered by the presence of attention. When you have attention, a phenomenal form is generated, which in turns triggers the *dharma* of reflexive awareness giving rise to a conscious momentary event. But there is only one of these for any conscious experience. There are further questions about the unity of consciousness at a time and over time. The Abhidharma, like all the other Buddhists, deny diachronic unity. Synchronic unity, they claim, can be explained by appeal to storehouse consciousness (for details about how such an account works see Chadha 2015).

Chalmers, however, suggests another justification which may help us to conceive protopanpsychist zombies: the nonquality/quality gap. But even this route is blocked by our Buddhist panprotopsychists: protophenomenal properties are not obviously nonqualitative properties. Protophenomenal properties range from what we straightforwardly conceive of as sensuous properties (e.g., redness) and feelings (e.g., pleasure, pain) or emotions (e.g., shame, fear) to mental faculties (e.g., attention), powers or dispositions (e.g., wisdom, concentration) and bodily properties (e.g., balance). On this Buddhist picture, I suggest, it is better to think in terms of degrees of consciousness, which in turn translates into degrees of qualitative or phenomenal awareness. The difference between phenomenal and protophenomenal properties is that of degree, not of kind. In other words, there is no unbridgeable nonquality/quality gap. Rather, what we have are the manifest and unmanifest states of conscious awareness. The unmanifest are not such that they cannot be consciously discerned but only that they are in the background of our manifest awareness. For example, in watching a beautiful sunset while walking on a rocky beach, I am manifestly aware of the beautiful red colour but there is a subtle unmanifest awareness of bodily balance in the background. The latter comes to the fore if I trip while walking. As already said, conscious awareness is an interplay of many different features and the intensity and character of the protophenomenal features determines the character of the phenomenal states.

This should give the reader an idea of the Abhidharma Buddhist view of consciousness and what it offers to the contemporary discussion of panpsychism. As I said before, my aim here has not been to present the Abhidharma Buddhist solution to various versions of the combination problem. I am only suggesting that these Buddhists ideas are worth exploring further if we are serious about investigating the solution space available to the panpsychists in addressing the combination problem.

Notes

1. *Skandhas* are aggregates of *dharmas* (1.7a-b).
2. The unconditioned *dharmas* are constituents of meditative states and liberation and are thus beyond the scope of this paper.
3. The idea that *dharmas* are tropes has been suggested in Mark Siderits (1997) and has also been developed by Ganeri (2001). Ganeri's focus is on an interpretation of Dignāga.
4. The talk about tropes is useful from a panpsychist point of view; it keeps the metaphysics of

substance and universals at an arm's length. Mental *dharmas* as protophenomenal parts are just elements of reality. The Abhidharma philosophers, and Buddhists more generally, do not want to ascribe to a substance metaphysics or a realism about universals.

5. *Ālaya-vijñāna* is posited in what Schmithausen (1987: 12, 18) calls the "initial passage" in the Basic Section of the *Yogācārabhūmi*. It is described as an unmanifest consciousness that persists within the material sense faculties during the highest meditative state (*nirodha samāpatti*, literally translated as the "attainment of extinction," signifying the extinction of perception and feeling). The later sections of the *Yogācārabhūmi* also offer other proofs for the existence of *ālaya-vijñāna*, some of which aim to provide a fix for the problem of *karmic* continuity. However, it is important to note that *ālaya-vijñāna* is one of a number of fixes proposed to deal with this problem.
6. Dharmakīrti's *Pramān˙avārttika* is referred to as PV *in accordance with the standard practice.*
7. Intention here is not to be understood as in contemporary Western philosophy as a state formed on the basis of belief-desire reasoning but rather as a *dharma*, a proto-conscious event that accumulates, organises, and rallies together other *dharmas* for the production of an object of a resulting conscious event.
8. As does Luke Roelofs in personal communication.

References

Baars, Bernard J. (1988). *A Cognitive Theory of Consciousness*. Cambridge, MA: Cambridge University Press.

Baumgarten, T. J., Schnitzler, A., and Lange, J. (2015). 'Beta Oscillations Define Discrete Perceptual Cycles in the Somatosensory Domain'. *Proceedings of the National Academy of Sciences of the United States of America*, 112 (39): 12187–92.

Chadha, M. (2015). 'The Problem of the Unity of Consciousness: A Buddhist Solution'. *Philosophy East and West*, 65 (3): 746–64.

Chalmers, David (1996). *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.

Chalmers, David (2015). 'Panpsychism and Panprotopsychism'. In T. Alter and Y. Nagasawa (eds.), *Consciousness in the Physical World: Essays on Russellian Monism*. Oxford: Oxford University Press.

Chalmers, David (2016). 'The Combination Problem for Panpsychism'. In G. Brüntrup and L. Jaskolla (eds.), *Panpsychism*. Oxford: Oxford University Press, pp. 229–48.

Dehaene, S., and Naccache, L. (2001). 'Towards a Cognitive Neuroscience of Consciousness: Basic Evidence and a Workspace Framework'. *Cognition*, 79: 1–37.

Dreyfus, G. (1997). *Recognizing Reality: Dharmakirti's Philosophy and Its Tibetan Interpretations*. Albany: State University of New York Press.

Dreyfus, G. (2011). 'Self and Subjectivity: A Middle Way Approach'. In M. Siderits, E. Thompson, and D. Zahavi (eds.), *Self, No-Self? Perspectives from Analytical, Phenomenological, and Indian Traditions*. Oxford: Oxford University Press, pp. 114–44.

Dreyfus, G., and Thompson, E. (2007). 'Asian Perspectives: Indian Theories of Mind'. In M. Moscovitch, E. Thompson, and P. Zezalo (eds.), *The Cambridge Handbook of*

Consciousness. Cambridge: Cambridge University Press.
Ganeri, J. (2001). *Philosophy in Classical India*. New York: Routledge.
Ganeri, J. (2012). *The Self: Naturalism, Consciousness and the First-Person Stance*. Oxford: Oxford University Press.
Garfield, Jay L. (2015). *Engaging Buddhism: Why It Matters to Philosophy*. New York: Oxford University Press.
Goodman, C. (2004). 'The *Treasury of Metaphysics* and the Physical World'. *Philosophical Quarterly*, 54 (216): 389–401.
Heim, M. (2014). *The Forerunner of All Things: Buddhaghosa on Mind, Intention, and Agency*. New York: Oxford University Press.
Jackson, Frank (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford and New York: Oxford University Press.
Lewis, David K. (1986). *On the Plurality of Worlds*. Oxford: Blackwell.
Maurin, A-S. (2014). 'Tropes'. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Fall edition. http://plato.stanford.edu/archives/fall2014/entries/tropes/.
Nagel, Thomas (1998). 'Conceiving the Impossible and the Mind-Body Problem'. *Philosophy*, 73 (285): 337–52.
Pruden, L. (1988). *Abhidharmakośabhāyam English Translation of Poussin, Louis de la Vallée., 1923–1931/1980. L'Abhidharmakośa de Vasubandhu*, 6 vols. Bruxelles: Institut Belge des Hautes Études Chinoises and Berkeley: Asian Humanities Press.
Ronkin, N. (2005). *Early Buddhist Metaphysics: The Making of a Philosophical Tradition* (Oxford Centre for Buddhist Studies Monograph Series). London and New York: Routledge-Curzon.
Seager, William, and Allen-Hermanson, Sean (2015). 'Panpsychism'. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Fall edition. http://plato.stanford.edu/archives/fall2015/entries/panpsychism/.
Schmithausen, L. (1987). *Ālaya-vijñāna: On the Origin and Early Development of a Central Concept in Yogacara Philosophy*. Tokyo: The International Institute for Buddhist Studies.
Siderits, Mark (1997). 'Buddhist reductionism'. *Philosophy East and West*, 47 (4), 455–478.
Siderits, Mark (2007). *Buddhism As Philosophy*. Indianapolis: Hackett.
Strawson, Galen (1999). 'The Self'. *Journal of Consciousness Studies*, 4 (5–6): 405–28. (Reprinted in S. Gallagher and J. Shear (eds.), *Models of the Self*. Thorverton: Imprint Academic, 1999.
Strawson, Galen (2006). *Consciousness and Its Place in Nature: Does Physicalism Entail Panpsychism?* Ed. A. Freeman. Exeter: Imprint Academic.
Strawson, P. F. (1985). *Skepticism and Naturalism: Some Varieties*. London: Methuen and New York: Columbia University Press.
Thompson, Evan (2014). *Waking, Dreaming, Being: Self and Consciousness in Neuroscience, Meditation, and Philosophy*. New York: Columbia University Press.
Varela, F., Lachaux, J-P., Rodriguez, E., and Martinerie, J. (2001). 'The Brainweb: Phase Synchronization and Large-Scale Integration'. *Nature Reviews – Neuroscience*, 2: 237.
Weatherson, B. (2015). 'Humean Supervenience'. In B. Loewer and J. Schaffer (eds.),