PANPSYCHISM

Contemporary Perspectives



Edited by GODEHARD BRÜNTRUP and LUDWIG JASKOLLA

Panpsychism

PHILOSOPHY OF MIND

Series Editor

David J. Chalmers , Australian National University and New York University

Self Expressions Owen Flanagan

Deconstructing the Mind Stephen Stich

The Conscious Mind David J. Chalmers

Minds and Bodies Colin McGinn

What's Within? Fiona Cowie

The Human Animal Eric T. Olson

Dreaming Souls Owen Flanagan

Consciousness and Cognition Michael Thau

Thinking Without Words José Luis Bermúdez

Identifying the Mind U.T. Place (author), George Graham, Elizabeth R. Valentine (editors)

Purple Haze Joseph Levine

Three Faces of Desire Timothy Schroeder

A Place for Consciousness Gregg Rosenberg

Ignorance and Imagination Daniel Stoljar

Simulating Minds Alvin I. Goldman

Gut Reactions Jesse J. Prinz

Phenomenal Concepts and Phenomenal Knowledge Torin Alter, Sven Walter (editors)

Beyond Reduction Steven Horst What Are We? Eric T. Olson

Supersizing the Mind Andy Clark

Perception, Hallucination, and Illusion William Fish

Cognitive Systems and the Extended Mind Robert D. Rupert

The Character of Consciousness David J. Chalmers

Perceiving the World Bence Nanay (editor)

The Contents of Visual Experience Susanna Siegel

The Senses Fiona Macpherson (editor)

Attention is Cognitive Unison Christopher Mole

Consciousness and the Prospects of Physicalism Derk Pereboom

Introspection and Consciousness Declan Smithies and Daniel Stoljar (editors)

The Conscious Brain Jesse J. Prinz

Decomposing the Will Andy Clark, Julian Kiverstein, and Tillmann Vierkant (editors)

Phenomenal Intentionality Uriah Kriegel (editor)

The Peripheral Mind István Aranyosi

The Innocent Eye Nico Orlandi

Does Perception Have Content? Edited by Berit Brogaard

The Varieties of Consciousness Uriah Kriegel

Panpsychism Godehard Brüntrup and Ludwig Jaskolla (editors)

Panpsychism

Contemporary Perspectives

EDITED BY GODEHARD BRÜNTRUP and LUDWIG JASKOLLA



OXFORD

UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide. Oxford is a registered trade mark of Oxford University Press in the UK and certain other countries.

Published in the United States of America by Oxford University Press 198 Madison Avenue, New York, NY 10016, United States of America.

© Oxford University Press 2017

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, without the prior permission in writing of Oxford University Press, or as expressly permitted by law, by license, or under terms agreed with the appropriate reproduction rights organization. Inquiries concerning reproduction outside the scope of the above should be sent to the Rights Department, Oxford University Press, at the address above.

> You must not circulate this work in any other form and you must impose this same condition on any acquirer.

Library of Congress Cataloging-in-Publication Data Names: Brüntrup, Godehard, editor. Title: Panpsychism : contemporary perspectives / edited by Godehard Brüntrup and Ludwig Jaskolla. Description: New York, NY : Oxford University Press, 2016. | Series: Philosophy of mind series | Includes index. Identifiers: LCCN 2016006039 | ISBN 9780199359943 (hardcover : alk. paper) Subjects: LCSH: Panpsychism. | Philosophy of mind. | Consciousness. Classification: LCC BD560.P34 2016 | DDC 141—dc23 LC record available at https://lccn.loc.gov/2016006039

1 3 5 7 9 8 6 4 2

Printed by Sheridan Books, Inc., United States of America

CONTENTS

List of Contributors vii

Introduction 1 godehard brüntrup and ludwig jaskolla

PART I THE LOGICAL PLACE OF PANPSYCHISM

- 1. Panpsychism and Panprotopsychism 19 DAVID J. CHALMERS
- 2. Emergent Panpsychism 48 GODEHARD BRÜNTRUP

PART II THE VARIETIES OF PANPSYCHISTIC ONTOLOGIES

- 3. Mind and Being: The Primacy of Panpsychism 75 GALEN STRAWSON
- 4. Panpsychism and Priority Cosmopsychism 113 YUJIN NAGASAWA AND KHAI WAGER
- 5. In Search of Mentons: Panpsychism, Physicalism, and the Missing Link 130

BERIT BROGAARD

6. Land Ho? We Are Close to a Synoptic Understanding of Consciousness 153

GREGG ROSENBERG

PART III PANPSYCHISM AND THE COMBINATION PROBLEM

7. The Combination Problem for Panpsychism 179

DAVID J. CHALMERS

8. What Combination Problem? 215

BARBARA GAIL MONTERO

9. Panpsychist Infusion 229

WILLIAM SEAGER

Panpsychism and Neutral Monism: How to Make up One's Mind 249

SAM COLEMAN

 The Phenomenal Bonding Solution to the Combination Problem 283

PHILIP GOFF

PART IV PANPSYCHISM AND ITS ALTERNATIVES

12. Mind Dust, Magic, or a Conceptual Gap Only? 305

BRIAN P. MCLAUGHLIN

13. Emergence and Panpsychism 334

ACHIM STEPHAN

14. Neutral Monism and Panpsychism 349

LEOPOLD STUBENBERG

15. Dualism and Panpsychism 369

CHARLES TALIAFERRO

16. Idealism and Panpsychism 387

UWE MEIXNER

Index 407

LIST OF CONTRIBUTORS

Berit Brogaard (University of Miami) **Godehard Brüntrup** (Munich School of Philosophy) **David J. Chalmers** (New York University) **Sam Coleman** (University of Hertfordshire) **Philip Goff** (Central European University) Ludwig Jaskolla (Munich School of Philosophy) Brian P. McLaughlin (Rutgers University) **Uwe Meixner** (University of Augsburg) **Barbara Gail Montero** (City University of New York) **Yujin Nagasawa** (University of Birmingham) **Gregg Rosenberg** (senior executive in private business) William Seager (University of Toronto Scarborough) Achim Stephan (Osnabrück University) Galen Strawson (University of Texas at Austin) **Leopold Stubenberg** (University of Notre Dame) **Charles Taliaferro** (St. Olaf College) **Khai Wager** (University of Birmingham)

Introduction

GODEHARD BRÜNTRUP AND LUDWIG JASKOLLA

καὶ ἐν τῷ ὅλῳ δή τινες αὐτὴν μεμῖχθαί φασιν

There are some, too, who say that soul is interfused throughout the universe.

—Aristotle on Thales, De Anima, 411a7

Panpsychism is as old as philosophy itself, a key idea in Western and Eastern philosophical traditions (Skrbina 2005). It comes in many forms, and the definition of the term varies. In his benchmark entry on *Panpsychism*, William Seager defined the term very broadly:

Panpsychism is the doctrine that mind is a fundamental feature of the world which exists throughout the universe. (Seager and Allen-Hermanson 2010)

It is thus distinguished from absolute idealism, according to which the world consists solely of minds and their activities. It is also distinguished from materialism, for which the world consists ultimately of mindless physical entities and their configurations. Finally, it is distinct from substance dualism, which assumes two categorically different realms of entities, mental and physical, which can possibly exist independently from each other. Panpsychists in contrast claim that mental being is a fundamental and ubiquitous feature of the universe but is not the only fundamental and ubiquitous feature of the universe. There are many ways to spell this out in a metaphysical system. The most common one in recent debate is the idea that physical structure as described in the formalized language of physics cannot by itself provide the ultimate grounding of reality but rather needs to be complemented by nonstructural intrinsic facts which escape the vocabulary of physics. These so-called quiddities are omnipresent in the cosmos, and at least some of them are metaphysically necessary to ground the emergence of higher levels of consciousness in the process of evolution. The astronomer Sir Arthur Eddington eloquently expressed this idea in *Space, Time, and Gravitation*:

Physics is the knowledge of structural form, and not knowledge of content. All through the physical world runs that unknown content, which must surely be the stuff of our consciousness. (Eddington 1920, 200).

Sir Bertrand Russell expressed a similar idea in *The Analysis of Matter*:

As regards the world in general, both physical and mental, everything we know of its intrinsic character is derived from the mental side, and almost everything we know of its causal laws is derived from the physical side. (Russell 1927, 402)

The most common response to panpsychism is an incredulous stare. This is perhaps caused by the idea that panpsychism entails the belief that mountains and rocks, molecules and elementary particles enjoy an anthropomorphically conceived conscious life. Most forms of panpsychism, however, distinguish between mere conglomerates like a rock formation and genuine individuals like animals and possibly elementary particles. Mental properties can only be attributed directly to genuine individuals. Could an elementary particle have some rudimentary form of mentality? Is this an idea divorced from reality taken right from armchair speculation in the 'metaphysical laboratory'? Not entirely. In contemporary quantum mechanics, one can easily find a variety of theories that ascribe some form of mentality to the quantum level. According to Henry Stapp's (2007) orthodox collapse interpretation, quantum mechanics is built upon psychophysical collapse events that mediate the emergence of actuality from potentiality and are intrinsically connected to mental properties. A somewhat similar view has been advanced by Roger Penrose and Stuart Hameroff (see Penrose 1996; Hameroff and Penrose 1996); according to them a collapse of the wave function is a moment of experience. According to Michael Epperson's ontological interpretation of the new standard 'decoherence' theory, physical and logical relations among quantum actualities drive the process of decoherence, that is, the logically conditioned actualization of mere potentialities. This account is based on a dipolar or dual-aspect metaphysics of nature in which physical and mental features of quantum events play a significant role (see Epperson 2012). Finally, in the nonstandard Bohmian interpretation of quantum mechanics, the individual particle is informed by the quantum potential which is 'active information' and is thus

Introduction

capable of representing information and action given, when that information within the limits provided by Schrödinger's equation (Bohm 1990; Bohm and Hiley 1993). This might well be seen as a very primitive case of mental representation (Pylkkänen 2006). Even at the level of the neurosciences, the explanatory plausibility of panpsychism has recently been explored (see Koch and Tononi 2015).

There is no shortage of attempts in the philosophy of physics to provide a space for mentality or something analogous to mentality at the fundamental level. The tendency to widen the scope of physics to allow it to contain the realm of the mental is, however, somewhat problematic. It seems that the scientific method confines physics to the mathematical-structural properties of the world. Any attempt to capture the phenomenal mind in this vocabulary will encounter all the conceptual problems (conceivability arguments, the knowledge argument or the argument from lack of analysis) that have been brought forward against physicalist reductionism.

This volume however focuses on the philosophical—strictly speaking metaphysical—arguments that have evolved from panpsychism. The attraction of panpsychism for philosophers of mind is that it seems to offer a world view which is capable of remaining within a broadly conceived naturalist framework whilst accounting for the emergence of a nonreductively conceived mental reality. The mind is neither explained by divine intervention nor by inexplicable strong emergence. In other words, it offers a genuine middle-way between physicalism and dualism. It also presents a welcome alternative to nonreductive physicalism which—according to Jaegwon Kim and others (e.g., Kim 2007)—might well be a conceptually unstable position.

So panpsychism, as treated here, is a move on the chessboard of metaphysics. Traditionally, two arguments have been advanced in favor of panpsychism. The first, called the genetic argument, is based on the philosophical principle 'ex nihilo, nihil fit.' If human consciousness is to evolve from a physical basis, then basic forms of mental being need to be present at the fundament of this evolutionary process. Both Thomas Nagel and William James have defended versions of this argument (Nagel 2012; James 1890). Second, the argument from intrinsic natures has been developed as a reaction to the claim that a complete ontology in terms of relations is possible. This view might be inconsistent based on model theoretic consequences of the Löwenheim-Skolem theorem: Relational patterns alone leave the intrinsic nature of concrete individuals underdetermined, and this is all we get from physics. Whilst the physical sciences tell us nothing of the intrinsic (more than merely structural) nature of matter, the intrinsic nature of matter is known in the case of human consciousness. As ontological monists, panpsychists claim that the intrinsic nature of matter in general is mental or proto-mental (see Eddington 1920).

This seems to be the most simple hypothesis consistent with the only data we have. This volume is, in part, an attempt to further develop those argumentative strategies.

Panpsychism is an option that deserves more attention in the future. We do not discuss the empirical question in this volume. It may well be the case that panpsychism will receive hardly any empirical justification. This leads some critics to the idea that panpsychism is useless speculation. The wellknown epistemic asymmetry between mental and physical properties, however, can account for this possible lack of empirical justification. We have no direct access to other phenomenal minds, not even that of fellow humans. If there are minimally structured and minimally rich traces of phenomenal features even at the very lowest levels of the universe—or maybe the whole of qualia-space somehow embedded as deep down as the Planck level, as Penrose and Hameroff (2014) suggest—then this will be so different from our conscious experience that it will seem completely alien to us. We could still posit something like such proto-mentality on theoretical grounds. This would then amount to an inference to the best explanation when trying to give a metaphysical account of the emergence of consciousness in evolution. In the words of William James:

If evolution is to work smoothly, consciousness in some shape must have been present at the very origin of things. (James 1890, 152)

Smoothness in this sense requires that mental properties of adjacent levels of nature are similar. But similarity is not a transitive relation. Mental properties at the fundamental level might well be wildly dissimilar from those found in living organisms. The highly speculative and abstract character of panpsychism has thus always to be kept in mind. Panpsychism is a possible move in the logical space of metaphysics, not a crude animalistic view of matter. Neither Leibniz's 'monads' (Rescher 1991) nor Whitehead's 'actual occasions' (Whitehead 1927) is endowed with something akin to human consciousness; they do have mental properties nevertheless. Both philosophers arrived at the need to posit the existence of proto-mentality in the process of trying to account for the existence of subjectivity, or "views from somewhere" in the world.

The idea of panpsychism was originally introduced into analytic philosophy of mind and metaphysics by Thomas Nagel in his *Mortal Questions* (see Nagel 1979, 181–95). In more contemporary debate the renaissance of panpsychism started, at least in a significant part, because of the alleged 'loophole' for the physicalist in David Chalmers's now classical zombie argument (see Chalmers 1996). Against the idea of zombies the standard scientific essentialist (or 'type-B materialist') argues that the zombie world is only conceivable but not metaphysically possible. The mistake, according to the type-B materialist, is that we are working with primary intensions when dealing with epistemic possibility (conceivability) and with secondary intensions when dealing with metaphysical possibility. In the first case, we consider a world as actual, in the latter case we consider a world as counterfactual (how things might have been but are not). If we consider Putnam's XYZ-world as actual, then 'water is not H_2O' is true. This is an instance of an epistemic possibility. If we rigidly hold the meaning of 'water' fixed by the actual world and counterfactually consider the XYZ-world, then 'water is not H_2O' turns out to be false. Then 'water is H_2O' expresses a Kripkean necessary truth.

Chalmers's argument is effective against this kind of countermove only if strengthened by the additional thesis that the epistemic possibility of the zombie world entails its metaphysical possibility. However, this in turn entails that the primary and secondary intensions coincide, which in the case of the description of all the physical facts is problematic. Physical properties are functionally defined. We can say that the primary intension of a physical predicate picks out whatever plays that functional role in a given world. We can also say that the secondary intension of a physical predicate is tied to the property playing that role in our world in such a way that in other possible worlds something else might play this role. In this case there would be possible worlds that satisfy the structural-mathematical description of our world in physical terms without being an exact duplicate of our world. The physical structure of those other worlds would be indistinguishable from our world, but the intrinsic natures grounding those relations would be different. Some of these worlds might well be zombie worlds. But then, zombie worlds are metaphysically possible. The type-B materialist can escape this move by assuming that there are indeed intrinsic properties of the physical in our world which are metaphysically necessary for the existence of phenomenal minds in our world. But then the primary and secondary intensions in the physical description of our world do not coincide. Chalmers's zombie argument against scientific essentialism seems to fail.

The price to be paid is a position that Chalmers called "Russellian monism" or "type-F monism," which he construed as a kind of panpsychism. But not all versions of Russellian Monism entail panpsychism. According to Chalmers, the different strains of type-F monism are made up of two alternatives: the first being 'neutral monism,' which claims that although there must be an intrinsic nature of matter, knowledge of most of the characteristics of this nature is barred from us (see Holman 2008). In contrast to this 'Kantian' move, on the panpsychistic version of type-F monism the intrinsic basis of the material world is experiential (or mental) in nature. Panpsychists have offered two arguments for the claim that their view is to be preferred to neutral monism: First, panpsychists advance the idea that if the intrinsic characteristics of nature are epistemically inaccessible to us, then neutral monism makes them even more of a mystery. Second, neutral monism does not help solve the genetic question; that is, how consciousness arises from a pure nonconscious basis, because the intrinsic basis of consciousness in neutral monism is defined as nonmental. Consequently, the genetic question is not solved but just pushed back to a level on which we have no positive epistemic grip.

Zombie worlds are then still conceivable but only when conceiving of reality exclusively using the primary intensions of concepts and thoughts. The zombie world is an indistinguishable copy of the structural-mathematical physical properties. It abstracts away from the intrinsic properties of the physical. We have now reached the starting point of this volume.

The Logical Place of Panpsychism. We open this volume with a discussion on the logical place of panpsychism in recent philosophy of mind. While Chalmers argues in favor of a constitutive version of panpsychism, Brüntrup defends the idea of a nonconstitutive panpsychism.

In his chapter "Panpsychism and Panprotopsychism," *David Chalmers* presents his Hegelian argument for panpsychism. It is inspired by Hegel's dialectical method in exploring the possibility of a conceptual middle-ground between materialism and dualism. It seeks out a 'synthesis' between these two antithetical positions.

Chalmers establishes this synthesis by dialectically elucidating the opposition of materialism and dualism, as well as their respective strengths and weaknesses:

Materialism is supported by causal arguments, which claim that causal explanations must be grounded in physical properties. If phenomenal properties are to be causally relevant, they have to be grounded in physical properties. This entails the truth of materialism.

Conceivability arguments undermine the truth of materialism: There is no strong modal entailment between physical and phenomenal properties; physicalism requires that physical facts necessitate all other facts.

Conversely, dualism is supported by conceivability arguments and heavily criticized by causal arguments.

Chalmers presents constitutive Russellian panpsychism as a promising synthesis in this dialectic. It preserves a robust naturalist outlook insofar as all higher-level entities can be fully understood as compositions of basic entities, thus excluding downward causation. Additionally, it can secure a causal

Introduction

role of the mental because phenomenal qualities ground all physical relations. The conceivability arguments against physicalism are correct insofar as they abstract from the intrinsic properties (quiddities) of the physical.

Russellian panpsychism can be contrasted with panprotopsychism: the thesis that the intrinsic properties of the physical microlevel are not phenomenal properties but are, nevertheless, able to collectively constitute phenomenal properties. Both positions are subsumed under the generic term 'constitutive Russellian monism.' The biggest challenge for this view is the so-called combination problem; Chalmers closes his chapter indicating that the Hegelian argument does not establish the truth of Russellian monism. Nevertheless, it clearly shows that type-F monism should be taken very seriously as means to explain consciousness.

The first part of *Godehard Brüntrup's* paper "Emergent Panpsychism" argues that panpsychism is a genuine and clearly delineated position in philosophy of mind. Like Chalmers, he situates panpsychism as an alternative to both dualism and physicalism. Additionally, he distinguishes it from both idealism and neutral monism.

Brüntrup focuses on Russellian panpsychism in its currently most discussed version. He defends a form of ontological monism, which takes the structural features of the world described by physics to be ontologically incomplete. They require a carrier, and phenomenal properties are the best candidate for being the ultimate carriers of physical structure. It is therefore a dual-aspect monism.

Brüntrup claims that nonconstitutive panpsychism has an edge over its constitutive cousin by accounting for the existence of macrosubjects via strong emergence. However, this is by the same token its greatest problem, since the avoidance of strong emergence seems to be the strongest motive for the endorsement of panpsychism.

To resolve this problem, he introduces a distinction between strong and superstrong emergence. Emergent dualism requires the latter and is thus open to an attack via the genetic argument for panpsychism. While Brüntrup's account allows for the emergence of new individuals, it does not involve inter-attribute emergence of the phenomenal originating from the nonphenomenal, and hence remains theoretically attractive.

The Varieties of Panpsychist Ontologies. In the second part, we selected contributions that explicate specific and surprisingly different accounts of panpsychism, ranging from atomism to holism. They are inspired by traditional philosophical idealism or the scientific search for ultimate particles.

In his contribution "Mind and Being: The Primacy of Panpsychism" *Galen Strawson* offers a twelve-word metaphysics. 'Stoff ist Kraft, Wesen ist Werden, Sein ist Sosein, Ansichsein ist Fürsichsein': concrete reality is force or energy and it is through-and-through processual (being is becoming); it admits no irreducible ontological distinction between substance and quality, and it consists of experience, experiencing, experientiality. He proposes that the first three claims are profoundly plausible metaphysical principles and that unprejudiced consideration of what we know about concrete reality obliges us to favor the fourth, that is, panpsychism or panexperientialism, above all other positive substantive metaphysical proposals.

The reason for favoring panpsychism over all other positive views about concrete reality is not just that panpsychism is the most ontologically parsimonious view (given that the existence of conscious experiencing is certain and that panpsychism doesn't posit the existence of any kind of stuff other than conscious experiencing). We also have to factor in the silence of physics: the point that physics with its numbers and equations is perfectly silent on the question of the intrinsic nonstructural nature of reality. It is an elementary mistake, given the silence of physics, to think that science or physics could ever support any belief in the existence of any nonexperiential reality, or ever give us reason to suppose that any concrete reality was more likely to be nonexperiential than experiential, and belief that the fundamental nature of concrete reality is nonexperiential gives rise to a problem that doesn't exist on the assumption that panpsychism is true—the problem of how experiential being can arise or emerge from fundamentally nonexperiential being. Strawson concludes with the suggestion that non-experiential concrete being may be metaphysically impossible.

Yujin Nagasawa and *Khai Wager* explicate a version of holistic panpsychism, which they call "priority cosmopsychism." Priority cosmopsychism is understood as a holistic alternative to classical micropanpsychism, which argues in favor of the ontological priority of the whole over its parts, combining it with the idea that this whole exemplifies phenomenal properties.

Priority cosmopsychism inherits the idea that there is—ontologically speaking—the whole and its respective parts from Schafferian priority monism. Despite this quasi-realistic concept of the respective parts, priority monists claim that the whole is fundamental in the sense that the parts are dependent on or derivative of the whole. To this is added a panpsychistic foundation: the cosmos as a whole is mentally—and in particular 'phenomenally'—propertied.

It is argued that priority cosmopsychism solves both the problem of infinite decomposition, because it refrains from the idea of ontological fundamentalism, and the combination problem, because there is no down-up combination constituting subjects of experience. Therefore, Nagasawa and Wager claim that priority cosmopsychism is theoretically more advantaged than panpsychism. They conclude their paper discussing some problems exclusive to priority cosmopsychism: the inexplicability of a cosmic consciousness, counterintuitiveness, and, last but not least, the seeming estrangement from science.

Introduction

Berit Brogaard proposes a version of constitutive panpsychism in her "In Search of Mentons: Panpsychism, Physicalism, and the Missing Link." Under the condition that one accepts Chalmers's arguments against type-B materialism, Brogaard argues that there are two contenders for explaining consciousness, the first being some version of strong emergence. She criticizes and finally rejects this kind of emergence in the first part of her chapter.

The second contender is panpsychism. Brogaard calls her proposal for a constitutive panpsychism 'the theory of mentons.' Following Searle, she argues that consciousness is best understood as a field-phenomenon, according to which the unified field of consciousness changes as informational content is added or deleted from the field. Consequently, her version of panpsychism is not a form of state-panpsychism in which individual particulars bear mental properties. She compares her account of mentons to gravitons, the hypothetical gauge bosons of a quantum theory of gravitation. Accordingly, Brogaard's mentons are construed as elementary particles that carry microexperiences.

Closing her chapter, Brogaard discusses different versions of the combination problem, which are answered by her account: the subject combination problem, the palette argument, and finally the revelation argument.

The section on the varieties of panpsychist ontologies is concluded by Gregg Rosenberg's paper "Land Ho? We Are Close to a Synoptic Understanding of Consciousness." He argues that we are closing in on a 'synoptic pyramid' which will allow us to understand the why's and wherefore's of consciousness at many levels of explanation; cognitive, biological, physical, and metaphysical. This new framework of understanding the general physical world basis of consciousness through the lens of integrated information invites metaphysical questions about why consciousness and information would be universally entwined. Because information is the product of causal constraint, an appropriate metaphysical view of causal constraint would shed light for philosophers on Tononi et al's findings. The theory of the causal nexus introduced in his 2004 A Place for Consciousness answers this call well, not only explaining why information and consciousness are entwined, but unpacking what is special about a causal nexus of mutual constraint between members. It explains specifically why integrated information and consciousness are entwined. Put together, the theory of natural individuals, the theory of integrated information, the theory of re-entrant processing and global workspace theory are promising to provide a comprehensive and completely compatible set of answers to the problems of consciousness at multiple levels of explanation, depth and detail.

The Combination Problem. The combination problem is probably the most powerful argument brought forth against historic and recent accounts of panpsychism. This is why we provide ample room for discussion of the

combination problem. Accordingly, the third part of the volume will explore the combination problem in detail and its different, complex dimensions.

David Chalmers's "The Combination Problem for Panpsychism" explores the conceptual landscape of the most important current critique of panpsychism—the combination problem.

Chalmers sets out from the 'locus classicus': William James's presentation of the combination problem in his 1890 *The Principles of Psychology*. He discerns three ways of formulating the problem, which revolve around three distinct characteristics of phenomenal states: the subject combination problem, the quality combination problem, and the structural combination problem.

Chalmers then turns to seven subversions of these three versions. The first four are connected to the subject combination problem:

- First, there is the classical Jamesian argument that there are no ontological aggregates. Since constitutive panpsychism presupposes aggregation, it must be false.
- Second, there is the subject-summing argument, which defends the idea that microsubjects never necessitate macrosubjects.
- Third, there is the conceivability argument, which defends the idea that a panpsychist zombie world is possible, thus showing that panpsychism is false.
- Fourth, there is the knowledge argument, which defends the idea that a Mary's black-and-white room scenario is possible, which would render panpsychism false.

The following two are connected to the quality combination problem:

- Fifth, there is the palette argument, which defends the idea that there are too few kinds of microexperiential properties in Russellian panpsychism to constitute a complex palette of macroexperiences.
- Sixth, there is the revelation argument, which defends the idea that since the nature of experience is revealed in introspection, and since we cannot introspect microexperiences, constitutive panpsychism cannot possibly be true.

The final explication is connected to the structural combination problem:

Seventh, there is the structural mismatch argument, which defends the idea that there is a structural mismatch between the macrophysical and the macrophenomenal, which is alleged to be inconsistent with the truth of constitutive Russellian panpsychism. Chalmers closes his paper by presenting various solutions and answers to these explications of the combination problem.

Barbara Gail Montero takes a radical approach to the combination problem. She argues in her contribution "What Combination Problem?" that the problem understood as a critique of panpsychism is ill-conceived because it searches for a solution to a question which the panpsychist should never have been asked: "I see the combination problem as iatrogenic: induced by philosophers in their attempts to cure panpsychism rather than following from panpsychism itself." Her central argument comes down to the following line of thought: If we refrain from thinking about the origin of higher-level forms of consciousness out of pools or mere groupings of proto-minds, but instead conceive of it along the lines of how higher-level individuals are generated by lower-level individuals in general, then we have no reason to believe that there is anything mysterious about the origins of higher forms of consciousness.

William Seager's solution to the combination problem is inspired by fusion accounts of emergence and builds upon ideas from William James and Alfred N. Whitehead. In his contribution "Panpsychist Infusion," he starts out from a twofold critique of the classical understanding of combination. He argues that we are indeed mistaken in thinking that combination is always in the "mechanical mode of causal composition." He takes quantum mechanics, as well as certain properties of black holes, to show that there are very good examples of combination that transcend this mechanical mode.

Seager's answer disentangles physical from mental combination. Fusion, he argues, is a psychological process. It is not the fusion of physical states into conscious states, but rather the fusion of mental states into new, more integrated mental states.

Sam Coleman's "Panpsychism and Neutral Monism: How to Make up One's Mind" tries to answer the combination problem by defending an alternative understanding of the proto-mental nature of basic particulars to that posited by panpsychism. He starts from a critical discussion of the alleged virtues of panpsychism, which motivates his own approach: Coleman describes panqualityism (his preferred alternative to panpsychism) as the thesis that the world's intrinsic nature consists of qualities that are 'unexperienced qualia.'

Coleman discusses how his approach fares in light of the various forms of the combination problem as introduced by Chalmers. In particular, he shows that panqualityism together with a HOT approach to consciousness helps answer the subject combination problem. He closes his paper with an evaluation of Chalmers's 'awareness zombies,' which have been introduced as an example against HOT-panqualityism.

Philip Goff's "The Phenomenal Bonding Solution to the Combination Problem" identifies certain theses which in conjunction entail the conceptual

core of the combination problem: Those theses-conceptual isolation of subjects, transparency conceivability principle, and phenomenal transparencylead to the following principle:"For any group of subjects, instantiating certain conscious states, it is possible that just subjects with those states exist in the absence of any further subject." The critic of panpsychism could conclude from this principle that lower-level conscious individuals cannot account for higher-level conscious individuals. Goff claims that this conclusion does not follow. This is because the above principle does not rule out the existence of a phenomenal bonding relation at the microlevel, that is to say a relation R such that when microlevel subjects bear R to each other, a new higher-level subject is necessarily brought into being. The above principle says nothing about relations, and so does not rule out this possibility. Certainly, these phenomenal bonding relations are not transparently known to us through introspection or in any other way; indeed, if we did have a grip on such relations we would never have found mental combination puzzling in the first place. But it is not clear there is any reason to assume that naturally evolved human beings have access to all features of reality.

Goff then considers the question of when and in what conditions subjects become phenomenally bonded and so form a further subject; conceiving of this question as one particular form of Peter van Inwagen's 'special composition question,' he explores various answers to it. Having rejected the alternatives, Goff is prepared to accept *unrestricted phenomenal composition*: subjects *always* combine to make further subjects. On the basis of this he tentatively proposes an empirical identity between the phenomenal bonding relation and the spatial relation.

Panpsychism and Its Alternatives. The final part of this volume is dedicated to the main current alternatives to panpsychism as means to explaining consciousness. The papers discuss the various arguments for and against panpsychism from the perspective of those alternatives.

Brian McLaughlin's "Mind Dust, Magic, or a Conceptual Gap?" opens this final part with an alternative classification of panpsychism's place in the philosophy of mind. He discusses a reformulation of Nagel's argument for panpsychism (which is a version of the genetic argument) in terms of qualia and identifies two conceptual gaps in it: First, the claim that the possession of nonphysical properties by the basic constituents of the world entails that these properties are qualia. Second, the claim that some basic constituents of the world have qualia entails that all basic constituents have qualia.

Trying to close these gaps, McLaughlin arrives at the following conclusion, which he takes to be the most coherent version of Nagel's argument:

At least some of the subatomic physical entities into which atoms are wholly decomposable have qualia; and at any level at which those subatomic entities are themselves wholly decomposable, they have some physical entities as constituents that have qualia.

McLaughlin then argues that both panpsychists and reductive physicalists will deny the second premise of the reformulation of Nagel's argument. It claims that qualia are never logically implied by nonqualitative properties. After discussing emergentism as an alternative, and after delineating the various problems of panpsychism, McLaughlin reports in his conclusion his wager for an adequate theory of consciousness: He would concede that there is an unbridgeable conceptual gap between qualia and other properties. It is incompatible with panprotopsychism. While both panpsychism and neurobiologicalism are compatible with this conceptual gap, McLaughlin believes, all things considered, the more plausible view is neurobiologicalism: this is an a posteriori identity thesis such that every quale is identical with a microphysical brain structure. There is an unbridgeable, conceptual gap between object and subject concepts that is not a metaphysical gap.

Achim Stephan's contribution on "Emergence and Panpsychism" addresses the question whether panpsychism is a superior alternative to emergentism. Stephan begins with the works of Galen Strawson and Godehard Brüntrup, who have both defended the idea that the emergence of experience from a wholly nonexperiential basis seems unintelligible. In his piece, Stephan delineates several forms of emergence and aims to show that panpsychism does not have an edge over at least some of them.

He begins with weak emergentism, which he takes to be defined by three characteristics: physical monism, the existence of systemic properties possessed only by the system but not its constituents, and synchronic nomological determination of systemic properties by its constituents. Via a definition of diachronic, evolutionary emergentism, Stephan supplements those basic characteristics by three additional features: novelty of the systemic properties, their structural unpredictability, and finally one of three forms of irreducibility.

Stephan then construes strong emergentism as a possible middle-path between clear-cut scientific explanation in terms of reduction and hard panpsychist realism about proto-mentality. Stephan closes his contribution by arguing that while emergentism accepts the irreducible nature of consciousness in the physical world, panpsychism faces the following problem: panpsychists are not able to provide reductive explanations for consciousness but propose the existence of proto-mental properties that are cognitively opaque to us and of which it is conceptually unclear how they form new unities of experience (the combination problem).

Leopold Stubenberg's contribution "Neutral Monism and Panpsychism" starts by defining neutral monism as the thesis that both the mental and the

physical are grounded in a neutral basis, which is neither mental nor physical. The neutral basis carves reality at its joints.

According to textbook definitions it seems reasonable to think that neutral monism is conceptually incompatible with panpsychism. But in recent debates, the definitions of panpsychism and neutral monism have become so fluid that they tend to collapse into each other.

Returning to Russell's original view, Stubenberg argues that the contrast between Russell's neutral monism and panpsychism is pronounced. Russell's own theory is very different from the various 'Russellianisms' currently discussed. Additionally, it seems able to incorporate many of those philosophical intuitions motivating panpsychism while avoiding many of its difficulties.

Charles Taliaferro's contribution "Dualism and Panpsychism" addresses and defends a theistic world view where there are categorical mind-physics and body-person distinctions.

His first section establishes Taliaferro's basic point: there is a substantial distinction between the mental and the physical grounded in an epistemic primacy of the mental over the physical. From this he concludes the genuine causal efficacy of the mental, which in turn substantiates Taliferro's claim that the mental poses an irreducible ontological category. Taliaferro thus begins with an epistemic point about the mental and experiential, and he then moves to a metaphysical case for mental causation.

In the second section, Taliaferro defends the idea that human persons are enduring entities irreducible to body or brain, and that such a substantial account of the human person leads toward dualism.

Both these forms of dualism are supported by the primacy of the mental, says Taliaferro. Here his endorsement of theism comes into play: There is no need to explain the emergence of the mind for the theist because God's mind is the basis for everything else. Theism can thus provide an explanation for the existence of the panpsychistic cosmos and for the different levels of consciousness pervading it.

Finally, *Uwe Meixner's* contribution is entitled "Idealism and Panpsychism." We have already seen that some panpsychists (e.g., Galen Strawson) have certain leanings toward idealism. According to Meixner, panpsychism is not a form of materialism. Thus, Meixner distinguishes four versions of panpsychism: dualistic atomistic panpsychism, dualistic holistic panpsychism, idealistic atomistic panpsychism, and finally idealistic holistic panpsychism.

All dualistic forms of panpsychism cannot account for the relation of the mental and the physical any better than straightforward emergent dualism. Atomistic forms of panpsychism suffer from problems already pointed out by William James.

Introduction

According to Meixner, the only appealing option for the panpsychist is idealistic holistic panpsychism. Meixner's approach is inspired by Edmund Husserl, George Berkeley, and Gottfried Wilhelm Leibniz. He argues that this approach neither denies the existence of the physical nor collapses into solipsism.

We would like to express our deep gratitude to the following scholars.

David Chalmers for encouraging us to pursue this project and for invaluable advice. David Chalmers's contribution "Panpsychism and Panprotopsychism," an earlier version of which was initially presented at our Munich 2011 conference, is a revised and adapted version of his 2013 Amherst Lecture in Philosophy. We are grateful for the possibility to print this revised version and would like to express our gratitude to the Amherst Lecture in Philosophy and David Chalmers because "Panpsychism and Panprotopsychism" is an indispensable reference point for all the other papers in this volume.

Philip Goff for important suggestions concerning our introduction. All the speakers of our 2011 conference "Emergence and Panpsychism" in Munich, Germany—Michael Blamauer, David Chalmers, Sam Coleman, Gernot Falkner, Philip Goff, Michael Hampe, Hong Yu Wong, Peter Jedlicka, Spyridon Koutroufinis, Riccardo Manzotti, Brian McLaughlin, Tobias Müller, Franz Riffert, Gregg Rosenberg, Christina Schneider, William Seager, David Skrbina, Kathrin Solhdju, Achim Stephan, and Galen Strawson.

The fact that both editors appear on the cover of this volume does not do justice to the many individuals, scholars, and institutions without whom this volume would not have been possible to realize. Our thanks go out to:

Peter Ohlin for his unwavering support during the long process of putting together a large volume like this. The anonymous referees for Oxford University Press for positive and constructive critique.

Liselotte and Ludwig Gierstl for their never-ending patience and philosophical advice in editing, proofreading, and supporting the development of this volume—Liselotte Gierstl for her indispensible support on the index. Katherine Dormandy and Stephen Henderson for their thorough proofreading of the non-native speaker contributions. Korbinian Friedl for helping with the bibliographies. Finally, we would like to thank the members of Godehard Brüntrup's research group in metaphysics and philosophy of mind.

This research was supported by Deutsche Forschungsgemeinschaft (DFG), the Volkswagen Stiftung, pro philosophia e.V., and Dr. h.c. mult. Erich J. Lejeune.

Today, we see research on panpsychism flourishing. The number of publications has significantly increased. It seemed to be about time to publish a major volume like this. We hope that this survey of important positions and developments will assist the reader in navigating this very fluid and active field in current philosophy of mind.

Munich, July 2015 Godehard Brüntrup and Ludwig Jaskolla

Bibliography

- A detailed and regularly updated bibliography on panpsychism in analytic philosophy of mind can be found at http://www.hfph.de/panpsychism
- Aristotle. De Anima. With Translation, Introduction, and Notes. Ed. R. D. Hicks. Cambridge: Cambridge University Press, 1907.
- Bohm, David. "A New Theory of the Relationship of Mind and Matter." *Philosophical Psychology* 3.2 (1990): 271–86.
- Bohm, David, and Basil Hiley. The Undivided Universe: An Ontological Interpretation of Quantum Mechanics. London: Routledge, 1993.
- Chalmers, David. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press, 1996.
- Eddington, Arthur. Space, Time and Gravitation an Outline of the General Relativity Theory. Cambridge: Cambridge University Press, 1920.
- Epperson, Michael. *Quantum Mechanics and the Philosophy of Alfred North Whitehead*. New York: Fordham University Press, 2012.
- Hameroff, Stuart, and Roger Penrose. "Conscious Events as Orchestrated Space-Time Selections. *Journal of Consciousness Studies* 3.1 (1996): 36–53.
- Hameroff, Stuart, and Roger Penrose. "Consciousness in the Universe: A Review of the 'Orch OR' Theory." *Physics of Life Review* 11.1 (2014): 39–78.
- Holman, Emmett. "Panpsychism, Physicalism, Neutral Monism, and the Russellian Theory of Mind." *Journal for Consciousness Studies* 15.5 (2008): 48–67.
- James, Williams. The Principles of Psychology. Cambridge: Harvard University Press, 1890.
- Koch, Christof, and Giulio Tononi. "Consciousness: Here, There, and Everywhere?" *Philosophical Transactions B* 370.1668 (2015). Web.
- Kim, Jaegwon. *Physicalism, or Something near Enough*. Princeton: Princeton University Press, 2007.
- Nagel, Thomas. Mortal Questions. Cambridge: Cambridge University Press, 1979.
- Nagel, Thomas. Mind and Cosmos. Why the Materialist Neo-Darwinian Conception of Nature Is Almost Certainly False. Oxford: Oxford University Press, 2012.
- Penrose, Roger. Shadows of the Mind: A Search for the Missing Science of Consciousness. Oxford: Oxford University Press, 1996.
- Pylkkänen, Paavo. Mind, Matter, and the Implicate Order. Berlin: Springer, 2006.
- Rescher, Nicholas. *G.W. Leibniz's Monadology. An Edition for Students*. Pittsburgh: University of Pittsburgh Press, 1991.
- Russell, Bertrand. The Analysis of Matter. London: Kegan Paul, 1927.
- Seager, William, and Sean Allen-Hermanson. "Panpsychism." Standford Encyclopedia of Philosophy. Ed. Edward Zalta. 2010. Web.
- Stapp, Henry. Mindful Universe. Quantum Mechanics and the Participating Observer. Berlin: Springer, 2007.
- Whitehead, Alfred North. Process and Reality: An Essay in Cosmology. New York: Free Press, 1929/1978.

PART I

THE LOGICAL PLACE OF PANPSYCHISM

1

Panpsychism and Panprotopsychism

DAVID J. CHALMERS

1.1 Introduction

Panpsychism, taken literally, is the doctrine that everything has a mind. In practice, people who call themselves panpsychists are not committed to as strong a doctrine. They are not committed to the thesis that the number two has a mind, or that the Eiffel Tower has a mind, or that the city of Canberra has a mind, even if they believe in the existence of numbers, towers, and cities.¹

Instead, we can understand panpsychism as the thesis that some fundamental physical entities have mental states. For example, if quarks or photons have mental states, that suffices for panpsychism to be true, even if rocks and numbers do not have mental states. Perhaps it would not suffice for just one photon to have mental states. The line here is blurry, but we can read the definition as requiring that all members of some fundamental physical types (all photons, for example) have mental states.

For present purposes, the relevant sorts of mental states are conscious experiences. I will understand panpsychism as the thesis that some fundamental physical entities are conscious: that is, that there is something it is like to be a quark or a photon or a member of some other fundamental physical type. This thesis is sometimes called *panexperientialism*, to distinguish it from other varieties of panpsychism (varieties on which the relevant entities are required to think or reason, for example), but I will simply call it panpsychism here.

Panpsychism is sometimes dismissed as a crazy view, but this reaction on its own is not a serious objection. While the view is counterintuitive to some, there is good reason to think that any view of consciousness must embrace some counterintuitive conclusions. Furthermore, intuitions about panpsychism seem to vary heavily with culture and with historical period. The view has a long history in both Eastern and Western philosophy, and many of the greatest philosophers have taken it seriously. It is true that we do not have much direct evidence for panpsychism, but we also do not have much direct evidence against it, given the difficulties of detecting the presence or absence of consciousness in other systems. And there are indirect reasons, of a broadly theoretical character, for taking the view seriously.

In this chapter I will present an argument for panpsychism. Like most philosophical arguments, this argument is not entirely conclusive, but I think it gives reason to take the view seriously. Speaking for myself, I am by no means confident that panpsychism is true, but I am also not confident that it is not true. This chapter presents what I take to be perhaps the best reason for believing panpsychism. A companion chapter, "The Combination Problem for Panpsychism" (Chalmers this volume), presents what I take to be the best reason for disbelieving panpsychism.

I call my argument the Hegelian argument for panpsychism. This is not because Hegel was a panpsychist. He seems to have been far from it, perhaps except insofar as he believed in a 'world-soul' (which suggests a sort of cosmopsychism, the view that the world as a whole is conscious). Rather, my argument takes the dialectical form often attributed to Hegel: the form of thesis, antithesis, synthesis.²

In my Hegelian argument, the thesis is materialism, the antithesis is dualism, and the synthesis is panpsychism. The argument for the thesis is the causal argument for materialism (and against dualism). The argument for the antithesis is the conceivability argument for dualism (and against materialism). Synthesized, these yield the Hegelian argument for panpsychism. In effect, the argument presents the two most powerful arguments for and against materialism and dualism, and motivates a certain sort of panpsychism as a view that captures the virtues of both views and the vices of neither.

It turns out that the Hegelian argument does not support only panpsychism. It also supports a certain sort of *panprotopsychism*: roughly, the view that fundamental entities are *proto-conscious*, that is, that they have certain special properties that are precursors to consciousness and that can collectively constitute consciousness in larger systems. Later in the chapter, I will examine the relative merits of panpsychism and panprotopsychism, and examine problems that arise for both.

1.2 Thesis and Antithesis: Materialism and Dualism

Our thesis is materialism (or physicalism): roughly, the thesis that everything is fundamentally physical. Our antithesis is dualism: roughly, the thesis that not everything is fundamentally physical, and the things that are not fundamentally physical are fundamentally mental. Our synthesis is panpsychism: very roughly, the thesis that everything is (or at least that some things are) fundamentally physical *and* fundamentally mental.

More specifically, we will be concerned with materialism and dualism about consciousness. Materialism about consciousness is the thesis that consciousness is fundamentally physical: that is, that truths about consciousness are grounded in the fundamental truths of a completed physics. Dualism about consciousness is the thesis that consciousness is not fundamentally physical: that is, that truths about consciousness are not grounded in the fundamental truths of a completed physics.

Grounding is a relation of metaphysical constitution. Truths about consciousness are grounded in physical truths if all truths in the first set obtain wholly in virtue of truths in the second set obtaining.³ The intuitive idea behind materialism is that physical truths somehow add up to and yield truths about consciousness. This requires at least that there is a metaphysically necessary connection between these truths, in that it is impossible for a world to be physically like ours without that world being phenomenally like ours. Intuitively, once God created the entities of physics, consciousness came along for free.

We will be especially concerned with microphysical properties and with phenomenal properties. Microphysical properties are the fundamental physical properties characterized by a completed physics. Microphysical entities are the fundamental physical entities characterized by that physics. (Despite the name, it is not definitionally required that these entities be small.) Microphysical truths are positive truths about the instantiation of microphysical properties by microphysical entities. Here a positive truth is intuitively a truth about the properties that an entity has, rather than those that it lacks (for more on this, see Chalmers 2012). Macrophysical properties (entities, truths) are those that are grounded in microphysical properties (entities, truths). For ease of discussion, I will use the word 'physical' to mean 'microphysical' throughout what follows, sometimes using 'microphysical' for explicitness.

Phenomenal (or experiential) properties are properties characterizing what it is to be a conscious subject. The most familiar phenomenal property is simply the property of phenomenal consciousness: An entity has this property when there is something it is like to be that entity. There are also many specific phenomenal properties, characterizing more specific conscious experiences. For example, phenomenal redness characterizes the distinct sort of conscious experience we have when we experience redness. An entity has the property of phenomenal redness when it has that sort of conscious experience. Phenomenal truths are positive truths about the distribution of phenomenal properties (i.e., truths about what it is like to be various entities). We can then say that materialism about consciousness is the thesis that all phenomenal truths are grounded in microphysical truths. Dualism about consciousness is the thesis that phenomenal truths are not all grounded in microphysical truths. In what follows, by 'materialism' and 'dualism' I mean materialism and dualism about consciousness.

We can put the conceivability argument against materialism (and for dualism) as follows. Here *P* is the conjunction of all microphysical truths about the universe, and *Q* is an arbitrary phenomenal truth (such as "I am conscious").

- (1) $P\&\sim Q$ is conceivable.
- (2) If $P\&\sim Q$ is conceivable, $P\&\sim Q$ is metaphysically possible.
- (3) If $P\&\sim Q$ is metaphysically possible, materialism is false.
- (4) Materialism is false.

Here we can say that a claim is conceivable when it is not ruled out a priori. So it is conceivable that there are mile-high unicycles, for example. A claim is metaphysically possible when it could have obtained: intuitively, when God could have created the world such that the claim would have been true. So it is plausibly metaphysically possible that there are mile-high unicycles.

Premise (1) here is supported by the conceivability of zombies: creatures microphysically identical to us without consciousness. Most people think that zombies do not actually exist, but there seems to be no a priori contradiction in the idea. Premise (2) is supported by general reasoning about the relationship between conceivability and possibility. The thesis needs to be refined to accommodate various counterexamples due to Kripke and others, but I will stay with the simple thesis here.⁴ Premise (3) is supported by the idea that if $P\&\sim Q$ is metaphysically possible, then P does not metaphysically necessitate Q, so Q is not grounded in P, since grounding plausibly requires metaphysical necessitation. Here the intuitive idea is that if God could have created a world microphysically identical to our world but without consciousness, then the presence of consciousness involves new fundamental properties over and above those of physics, so materialism is false.

The conceivability argument is an epistemic argument against materialism, starting with an epistemological premise and proceeding to a metaphysical conclusion. There are other closely related epistemic arguments. These include the knowledge argument, which starts from the premise that Q is not deducible from P and concludes that it is not grounded in P; the explanatory argument, which starts from the premise that there is an explanatory gap between P and Q and concludes that there is an ontological gap; and the structure-dynamics argument, which starts from the premise that P can be analyzed in terms of

structure and dynamics while Q cannot and concludes that Q is not grounded in *P*. Much of what I say will apply to all these arguments, but I will focus on the conceivability argument here.

Materialists do not just curl up and die when confronted with the conceivability argument and its cousins. Type-A materialists reject the epistemic premise, holding for example that zombies are not conceivable. Type-B materialists reject the step from an epistemic premise to an ontological conclusion, holding for example that conceivability does not entail possibility. Still, there are significant costs to both of these views. Type-A materialism seems to require something akin to an analytic functionalist view of consciousness, which most philosophers find too deflationary to be plausible. Type-B materialism seems to require a sort of brute necessity that is not found elsewhere and that is hard to justify. Of course some philosophers find these costs worth paying, or deny that these are costs. Still, I think that the argument makes at least a prima facie case against materialism.

That said, many materialists think that the conceivability argument against materialism (and for dualism) is countered by the causal argument against dualism (and for materialism). This argument runs as follows:

- (1) Phenomenal properties are causally relevant to physical events.
- (2) Every caused physical event has a full causal explanation in physical terms.
- (3) If every caused physical event has a full causal explanation in physical terms, every property causally relevant to the physical is itself grounded in physical properties.
- (4) If phenomenal properties are grounded in physical properties, materialism is true.
- (5) Materialism is true.

Here we can say that a property is causally relevant to an event when instantiations of that property are invoked in a correct causal explanation of that event. For example, the high temperatures in Victoria were causally relevant to the Victorian bushfires. A full causal explanation of an event is one that characterizes *sufficient* causes of the event: causes that guarantee that the event will occur, at least given background laws of nature.

Premise (1) is supported by intuitive observation. My being in pain seems to cause my arm to move. If things are as they seem here, then the pain will also be causally relevant to the motion of various particles in my body. Premise (2) follows from a widely held view about the character of physics: Physics is causally closed, in that there are no gaps in physical explanations of physical events. Premise (3) is a rejection of a certain sort of overdetermination. Given a full microphysical causal explanation of physical events, other causal explanations are possible only when the factors involved in the latter are grounded in the factors involved in the former (as when we explain the motion of a billiard ball both in terms of another ball and in terms of the particles that make it up).⁵ Any putative causal explanation that was not grounded in this way would involve causal overdetermination by independent events. Systematic overdetermination of this sort is widely rejected. Premise (4) is true by definition.

Dualists do not just curl up and die when presented with the causal argument. Epiphenomenalists reject premise (1), holding that the claim that consciousness causes behavior is just an intuition and can be rejected. Interactionists reject premise (2), holding that physics leaves room for (and perhaps is positively encouraging to) causal gaps that consciousness might fill. Still, there are costs to both of these views. Epiphenomenalism is at least inelegant and requires special coincidences between conscious experiences and macrophysical events (utterances about consciousness, for example) that seem to reflect them. Interactionism requires a view of physics that would be widely rejected by most physicists, and that involves a large bet on the future of physics. Again, some dualists (including me in some moods) deny that these are costs or hold that the costs are worth paying. Still, I think there is at least a prima facie case against dualism here.

So we have a standoff. On the face of it, the conceivability argument refutes materialism and establishes dualism, and the causal argument refutes dualism and establishes materialism. It is time for a Hegelian synthesis.

1.3 Synthesis: Panpsychism

Panpsychism, once again, is the thesis that some microphysical entities are conscious. For our purposes, it is useful to distinguish various more fine-grained varieties of panpsychism. To do this, we can first introduce some terminology.

Let us say that *macroexperience* is the sort of conscious experience had by human beings and other macroscopic entities (i.e., entities that are not fundamental physical entities). Macroexperience involves the instantiation of *macrophenomenal properties*: properties characterizing what it is like to be humans and other macroscopic entities. Let us say that *microexperience* is the sort of conscious experience had by microphysical entities. Microexperience involves the instantiation of microphenomenal properties: properties characterizing what it is like to be microphysical entities.

If panpsychism is correct, there is microexperience and there are microphenomenal properties. We are not in a position to say much about what microexperience is like. I think we can be confident that it is very different from human experience, however. It is almost certainly much simpler than human experience. In the way that an experience of redness is much simpler than a stream of conscious thought, we can expect a quark's experience to be much simpler than an experience of redness. To get far beyond generalities like this concerning microexperience, we would need a proper panpsychist theory of consciousness, which we are currently lacking.

Constitutive panpsychism is the thesis that macroexperience is (wholly or partially) grounded in microexperience. More or less equivalently, it is the thesis that macroexperience is constituted by microexperience, or realized by microexperience. On this view, macrophenomenal truths obtain in virtue of microphenomenal truths, in roughly the same sense in which materialists hold that macrophenomenal truths obtain in virtue of microphysical truths. To put things intuitively, constitutive panpsychism holds that microexperiences somehow add up to yield macroexperience. The view can allow that macroexperience is not wholly grounded in microexperience: For example, it might be grounded in microexperience along with certain further structural or functional properties.

Panpsychists need not be constitutive panpsychists. There is also nonconstitutive panpsychism, which holds that there is microexperience and macroexperience, but the microexperience does not ground the macroexperience. Nonconstitutive panpsychists will typically be *emergent panpsychists*, holding that macroexperience is strongly emergent from microexperience and/or from microphysics. One sort of emergent panpsychist might hold that there are contingent laws of nature that specify when certain microexperiences give rise to certain macroexperiences. Another might hold that there are laws of nature connecting microphysical properties to microphenomenal properties and macrophysical properties to macrophenomenal properties, without there being any constitutive connection between microphenomenal and macrophenomenal. Still, as we will see, nonconstitutive panpsychism inherits many of the problems of dualism, while it is constitutive panpsychism that offers hope for a Hegelian synthesis. So it is this view that I will focus on.

Like materialism, constitutive panpsychism comes in type-A and type-B varieties. Type-A constitutive panpsychism holds that there is an a priori entailment from microphenomenal truths to macrophenomenal truths, while type-B constitutive panpsychism holds that there is an a posteriori necessary entailment from microphenomenal truths to macrophenomenal truths. The type-B view inherits many of the problems of type-B materialism, so it is the type-A view that offers special hope for a Hegelian synthesis. When I talk of constitutive panpsychism, it will usually be the type-A version that I have especially in mind. Another important variety of panpsychism is *Russellian panpsychism*. This view takes its name from Russell's insight, in *The Analysis of Matter* (Russell 1927) and other works, that physics reveals the relational structure of matter but not its intrinsic nature. According to this view, classical physics tells us a lot about what mass does—it resists acceleration, attracts other masses, and so on—but it tells us nothing about what mass intrinsically is. We might say that physics tells us what the mass role is, but it does not tell us what property plays this role.

Here we can say that *quiddities* are the fundamental categorical properties that play the fundamental roles specified in physics. Alternatively, we can say that quiddities are the categorical bases of the microphysical dispositions characterized in physics. We can stipulate in addition that quiddities are distinct from the roles or the dispositions themselves. A view on which there are only role or dispositional properties, and no distinct properties playing those roles or serving as the basis for the dispositions, is a view on which there are no quiddities.

It is not obvious that there must be quiddities. There are respectable structuralist or dispositionalist views of physics on which physics involves just structure or dispositions all the way down. Still, many find these views objectionable, because they seem to yield a world devoid of substance or qualities— Russell said that on views like these "all the things in the world will merely be each others' washing" (Russell 1927, 325). And whether or not one accepts these objections, it is certainly not obvious that there are no quiddities. On the face of it, a worldview that postulates quiddities is perfectly coherent, and there is little clear evidence against it.

Russellian panpsychism is the view that some quiddities are microphenomenal properties. This view requires that there are quiddities—distinct properties that play the mass role, the charge role, and so on—and that at least some of these quiddities are phenomenal. For example, perhaps the property that plays the mass role is a certain phenomenal property. (Or better, as mass is really a quantity: The quantity that plays the mass role is a certain phenomenal quantity.) The Russellian panpsychist addresses two metaphysical problems what is the place of phenomenal properties in nature, and what are the intrinsic properties underlying physical structure?—and in effect answers both of them at once. Fundamental phenomenal properties play fundamental microphysical roles and underlie fundamental microphysical structure.

Panpsychists need not be Russellian panpsychists. There is also non-Russellian panpsychism, according to which there are microphenomenal properties that do not play microphysical roles. Perhaps there are numerous microphenomenal properties quite distinct from the properties involved in the microphysical network, for example. Still, non-Russellian panpsychism faces obvious problems with mental causation, while Russellian panpsychism offers hope for a Hegelian synthesis. So it is this view that I will focus on.

In particular, I will focus on constitutive Russellian panpsychism. On this view, microphenomenal properties serve as quiddities, playing the roles associated with microphysical properties, and also serve as the grounds for macrophenomenal properties. That is, microexperience constitutes macroexperience while also playing microphysical roles. On this view, one could think of the world as fundamentally consisting in fundamental entities bearing fundamental microphenomenal properties, where these microphenomenal properties are connected to each other (and perhaps to other quiddities) by fundamental laws with the structure that the laws of physics describe. All this microphenomenal structure also serves to constitute the macrophenomenal realm, just as microphysical structure serves to constitute the macrophysical realm.

I think that constitutive Russellian panpsychism is perhaps the most important form of panpsychism, precisely because it is this form that promises to avoid the problems of physicalism and dualism and to serve as a Hegelian synthesis. In particular, one can argue that this view avoids both the conceivability argument against physicalism and the causal argument against dualism.

To assess this matter, we first need to assess a delicate question: Is constitutive Russellian panpsychism a form of materialism, a form of dualism, or neither? This question turns on the answer to another delicate question: Are quiddities physical properties? If quiddities are physical properties, then constitutive Russellian panpsychism entails that microphenomenal properties are physical properties, and that macrophenomenal properties are constituted by physical properties, so that materialism is true. If quiddities are not physical properties, however, then macrophenomenal properties will be constituted by nonphysical properties, and a form of dualism will be true.

To answer this question, it is useful to make a distinction. We can say that *narrowly physical* properties are microphysical role properties, such as the dispositional property associated with having a certain mass, or the second-order property of having a property that plays the mass role.⁶ We can say that *broadly physical* properties are physical role properties along with any properties that realize the relevant roles: categorical bases for the mass dispositions, first-order properties that play the mass role.

In effect, narrowly physical properties include structural properties of microphysical entities but exclude quiddities, while broadly physical properties include both structural properties and quiddities. Here a structural property is one that can be fully characterized using structural concepts alone, which I take to include logical, mathematical, and nomic concepts, perhaps along with spatiotemporal concepts (see Chalmers 2003 and 2012 for much more discussion). If one uses a Ramsey sentence to characterize fundamental
physics, it is plausible that one can do so using structural concepts alone. At the same time, if there are quiddities, it is plausible that they (like phenomenal properties) cannot be fully characterized in structural terms.

We can then say that quiddities are not narrowly physical, but they are broadly physical. There is more to say here, particularly concerning just how we should construe the relation between quiddities and ordinary physical properties such as mass, but I will leave this for the next section.

With this distinction made, the question of whether quiddities are physical properties becomes something of a verbal question. One can use the term 'physical' to cover only narrowly physical properties or to cover broadly physical properties, and the choice between these usages is a verbal matter. Some may think that there is a stronger case for one usage or the other, but little of substance turns on this.

The same applies to the question of whether constitutive Russellian panpsychism is physicalism. We can distinguish narrow physicalism, which holds that phenomenal truths are grounded in narrowly physical truths, from broad physicalism, which holds that phenomenal truths are grounded in broadly physical truths. Narrow physicalism entails broad physicalism, but broad physicalism may not entail narrow physicalism. In particular, constitutive Russellian panpsychism is incompatible with narrow physicalism, but it is a form of broad physicalism. Once again, any dispute over whether narrow or broad physicalism is really physicalism will be something of a verbal dispute. Instead, constitutive Russellian panpsychism falls into a penumbral area that might be counted either way. This is a promising area for a Hegelian synthesis.

How does constitutive Russellian panpsychism fare with respect to the conceivability argument against physicalism? Once we have the distinction between narrowly and broadly physical truths in place, we can distinguish two different versions of the argument. One version construes P as the conjunction of all positive narrowly physical truths, takes as a premise that the corresponding version of P&~Q is conceivable, and concludes that narrow physicalism is false. The other does the same for broadly physical truths and broad physicalism.

To assess these arguments, we can distinguish two different sorts of zombies: narrowly physical duplicates of us without consciousness, and broadly physical duplicates of us without consciousness. We can call the first group *structural zombies*, since they duplicate just our relational physical structure. We can call the second group *categorical zombies*, since they also duplicate the underlying categorical properties.⁷

It is plausible that when we typically conceive of zombies, we are really conceiving of structural zombies. We hold physical structure fixed, but we do not make any effort to hold quiddities fixed, since we have no idea what the quiddities are. This standard zombie intuition provides good reason to think that structural zombies are conceivable, but little reason to think that categorical zombies are conceivable. If this is right, adding the conceivabilitypossibility premise at best establishes the possibility of structural zombies but not of categorical zombies. This is a happy result for (type-A) constitutive Russellian panpsychists, who hold that categorical zombies are not conceivable and not possible.

The upshot of this is that the standard considerations about conceivability can be used at most to undermine narrow physicalism but not broad physicalism. So these considerations have no force against constitutive Russellian panpsychism, which is a version of the latter but not the former. It follows that this view evades at least one horn of the Hegelian dilemma.

What about the other horn: the causal argument against dualism? Here it is useful to first reflect on the causal role of experience under constitutive Russellian panpsychism. According to Russellian panpsychism, microphenomenal properties certainly play a causal role in physics. They are the properties that play the most fundamental causal roles in physics: the mass role, the charge role, and so on. A microphenomenal property that plays the mass role is causally responsible for attracting other entities, and so on. This causation does not involve any violation of the laws of physics. Instead, this sort of causation underlies the laws of physics.

At the same time, constitutive panpsychism allows that macroexperience can inherit causal relevance from microexperience. This is an instance of the general claim that constituted properties can inherit causal relevance from constituting properties. For example, a billiard ball can inherit causal relevance from that of the particles that make it up. I think this is the lesson of much recent discussion of causal exclusion between the microscopic and macroscopic levels: when entities at this level are constitutively connected, there need be no causal exclusion. The moral that applies to the microphysical and the macrophysical also applies to the microphenomenal and the macrophenomenal, if they are constitutively connected.

It follows that constitutive Russellian panpsychism is compatible with a robust causal role for both microexperience and macroexperience. Given that microexperience is causally relevant (as Russellian panpsychism suggests), and that microexperience constitutes macroexperience (as constitutive panpsychism suggests), we can expect that macroexperience will be causally relevant, too.

What of the causal argument? Here again we need to distinguish versions of the argument. One version of the argument invokes the causal closure of the broadly physical to argue that phenomenal properties are grounded in broadly physical properties. The premises of this version of the argument are all plausible, and the constitutive Russellian panpsychist can happily accept its conclusion. Another version invokes the causal closure of the narrowly physical to argue that phenomenal properties are grounded in narrowly physical properties. Here the constitutive Russellian panpsychist must reject the conclusion, but fortunately they can easily reject premise (2). A full causal explanation of narrowly physical events will involve broadly physical properties; a causal explanation wholly in terms of narrowly physical properties is incomplete. This is to say that on a view where there are quiddities, the broadly physical domain may be causally closed, but the narrowly physical domain will not be.⁸

The upshot is that the causal argument can be used at best to establish broad physicalism and not narrow physicalism. This is once again a happy result for the constitutive Russellian panpsychism, as it is a version of the former but not the latter. So this view evades the second horn of our Hegelian dilemma.

We can combine our analysis of the two arguments as follows. The conceivability argument refutes narrow physicalism but is compatible with broad physicalism. The causal argument establishes broad physicalism but does not establish narrow physicalism. When these arguments are put together, they yield the Hegelian argument for the conjunction of broad physicalism with the denial of narrow physicalism. This is the ground occupied by constitutive Russellian panpsychism.⁹

It is worth noting that non-constitutive and non-Russellian panpsychism do not evade the Hegelian dilemma. Both of these views are incompatible with broad physicalism, and so are vulnerable to the causal argument for broad physicalism. On nonconstitutive panpsychism, even if microexperience is causally relevant, macroexperience will lie outside the broad physical network, so it will lead to epiphenomenalism, interactionism, or overdetermination. On non-Russellian panpsychism, it is hard to see how even microphenomenal properties can be causally relevant, and the same trilemma ensues. Among versions of panpsychism, only constitutive Russellian panpsychism promises to serve as a Hegelian synthesis.

1.4 Antithesis: Panprotopsychism

It is a familiar point in the pseudo-Hegelian dialectic that every synthesis is confronted by a new antithesis and followed by a new synthesis. Our Hegelian synthesis above is panpsychism. But it turns out that another view can also escape the original Hegelian dilemma: constitutive Russellian panprotopsychism. Recall that panprotopsychism is the view that fundamental physical entities are protoconscious. In more detail, let us say that *protophenomenal* properties are special properties that are not phenomenal (there is nothing it is like to have a single protophenomenal property) but that can collectively constitute phenomenal properties, perhaps when arranged in the right structure. Panprotopsychism is then the view that some fundamental physical entities have protophenomenal properties.

One might worry that any nonpanpsychist materialism will be a form of panprotopsychism. After all, nonpanpsychist materialism entails that microphysical properties are not phenomenal properties and that they collectively constitute phenomenal properties. This is an undesirable result. The thought behind panprotopsychism is that protophenomenal properties are special properties with an especially close connection to phenomenal properties. To handle this, one can unpack the appeal to specialness in the definition by requiring that (i) protophenomenal properties are distinct from structural properties and (ii) that there is an a priori entailment from truths about protophenomenal properties (perhaps along with structural properties) to truths about the phenomenal properties that they constitute. This excludes ordinary type-A materialism (which grounds phenomenal properties in structural properties) and type-B materialism (which invokes an a posteriori necessary connection). From now on I will understand protophenomenal properties this way and will understand panprotopsychism accordingly.¹⁰

I have occasionally heard it said that panprotopsychism can be dismissed out of hand for the same reason as materialism. According to this objection, the epistemic arguments against materialism all turn on there being a fundamental epistemic (and therefore ontological) gap between the nonphenomenal and the phenomenal: There is no a priori entailment from nonphenomenal truths to phenomenal truths. If this were right, the gap would also refute panprotopsychism. I do not think that this is right, however. The epistemic arguments all turn on a more specific gap between the physical and the phenomenal, ultimately arising from a gap between the structural (or the structural/ dynamical) and the phenomenal. We have principled reasons to think that phenomenal truths cannot be wholly grounded in structural truths. But we have no correspondingly good reason to think that phenomenal truths cannot be wholly grounded in nonphenomenal (and nonstructural) truths, as panprotopsychism suggests.

It is true that we do not have much idea of what protophenomenal properties are like. For now they are characterized schematically, in terms of their relation to phenomenal properties. A fuller account will have to wait for a full panprotopsychist theory, though I will speculate about one sort of protophenomenal property toward the end of this chapter. But our ignorance about protophenomenal properties should not be mistaken for an objection to the truth of panprotopsychism.

Constitutive panprotopsychism is roughly the thesis that macroexperience is grounded in the protophenomenal properties of microphysical entities. That is, all phenomenal truths are grounded in protophenomenal truths concerning these entities. As before, constitutive panprotopsychism could in principle come in type-A and type-B varieties, but the definition of specialness above in effect restricts it to the type-A version (a priori entailment from protophenomenal truths to macrophenomenal truths), which is in any case the relevant version for our purposes. *Russellian* panprotopsychism is the thesis that some quiddities are protophenomenal properties. For example, perhaps protophenomenal properties play the mass role or the charge role.

Nonconstitutive and non-Russellian panprotopsychism are coherent theses as protophenomenal properties are defined above (at least if we set aside the specialness clause): perhaps protophenomenal properties only constitute some macroexperiences, and perhaps they do not serve as quiddities. As with panpsychism, however, the Hegelian motivations for panprotopsychism strongly favor (type-A) constitutive Russellian panprotopsychism, so it is this view on which I will concentrate.

Constitutive Russellian panprotopsychism, like constitutive Russellian panpsychism, is a form of broad physicalism without narrow physicalism. It therefore escapes the Hegelian dilemma in just the same way. Constitutive Russellian panpsychists will reply to the conceivability argument by saying that structural zombies are conceivable but categorical zombies are not. They will reply to the causal argument by saying that fundamental protophenomenal properties are causally relevant in virtue of playing microphysical roles, and that macrophenomenal properties inherit causal relevance from protophenomenal properties in virtue of being grounded in them. In this way it slips through the horns of the Hegelian dilemma.

1.5 Synthesis: Russellian Monism

Given panpsychism as thesis and panprotopsychism as antithesis, there is a natural synthesis that subsumes them both. This synthesis is Russellian monism.¹¹ We can understand Russellian monism as the conjunction of broad physicalism with the denial of narrow physicalism. On this view, structural properties in physics do not constitute consciousness, but quiddities (perhaps along with structure) constitute consciousness. The view is Russellian because of the appeal to quiddities and their connection to mentality. It is a sort of monism because the world on this view consists in quiddities connected by laws of nature.

It is easy to see that both constitutive Russellian panpsychism and constitutive Russellian panprotopsychism are forms of Russellian monism. In fact, Russellian monism is equivalent to the disjunction of the two. According to Russellian monism, all conscious experience is grounded in structure plus quiddities but not in structure alone. Given the definition of protophenomenal properties above, this thesis is equivalent to the thesis that some quiddities are either phenomenal or protophenomenal, as the Russellian views hold, and that these quiddities along with structure ground all conscious experience, as the constitutive views hold.

Is Russellian monism a form of physicalism, dualism, or something else? As before, this is a largely verbal question that we need not settle. We could say that it is a form of broad physicalism but not narrow physicalism, and leave it at that. Still, it is interesting to look more closely at the question of whether, on a Russellian monist view, (proto)phenomenal properties (i.e., phenomenal or protophenomenal properties) are physical properties. There are a number of different options available here, depending on what one counts as a physical property, and how one construes the semantics of physical terms such as 'mass.' Each of these options leads to a subtly different way of characterizing Russellian monism. The following discussion may be of most interest to aficionados of this topic; others can skip it without much loss.

An initial question is whether physical properties are restricted to the properties invoked by physical theory (space, time, mass, charge, and so on), perhaps along with those properties grounded in them. These are the properties that Stoljar calls the t-physical properties (for theory-physical) and that Strawson (2006) calls 'physics-al' properties. It is most common to restrict physical properties in this sense, but one can also invoke expanded senses of the term, such as my notion of a broadly physical property, or Stoljar's notion of an o-physical property, or Strawson's notion of a physical property which appears to subsume all natural properties. Given such an expanded sense, then even if quiddities are not t-physical properties, they may count as physical in the expanded sense. The resulting position might be seen as *expansionary Russellian physicalism*, with (proto)phenomenal properties counting as physical properties in an expanded sense.¹²

In what follows, I will make the more common assumption that physical properties are restricted to t-physical properties: perhaps space, time, mass, charge, and so on. To assess the status of Russellian monism, we can then ask: What is the relationship between (proto)phenomenal properties and physical properties such as mass? This depends on just how terms such as 'mass' function. On one view, 'mass' refers to the property that actually plays the mass role. So insofar as there is a quiddity that actually plays the mass role, that quiddity is identical to mass. The corresponding version of Russellian monism is the *Russellian identity theory*, because it holds that (proto)phenomenal properties are identical to physical properties such as mass. As Grover Maxwell observes, this is a sort of inversion of the more familiar identity theory due to Smart and others (Maxwell 1978; Smart 1959). The familiar identity theory offers a topic-neutral analysis of mental expressions, where 'pain' refers to whatever plays the pain role, and then holds that these have physical referents, with C-fiber firing playing the pain role. The Russellian identity theory instead offers a topic-neutral analysis of physical expressions, where 'mass' refers to whatever plays the mass role, and then holds that these have mental or proto-mental referents, with (proto)phenomenal quiddities playing the mass role.¹³

On another view, 'mass' refers to the second-order functional property of having a property that plays the mass role. On this view, mass is not identical to the quiddity that plays the mass role, but we might say that mass is realized by that quiddity. A closely related view holds that 'mass' refers to a dispositional property which is realized by the quiddity that serves as its categorical basis. The corresponding version of Russellian monism is the *Russellian realization theory*, since it holds that physical properties such as mass are realized by (proto)phenomenal properties. Russellian realization theory can be seen as an inversion of the familiar functionalist realization theory, on which mental properties are second-order functional properties (pain is the property of having a property that plays the pain role) and on which these properties are realized by physical properties.

On the Russellian realization theory, quiddities are not themselves tphysical properties (at least if we assume that realizing properties are distinct from the properties they realize). So the Russellian realization theory is not a version of physicalism, assuming as above that only t-physical properties are physical properties. Instead, physical properties are themselves realized by and grounded in the (proto)phenomenal properties that serve as quiddities. The panpsychist version of this view can be seen as a form of *Russellian idealism*, with fundamental phenomenal properties serving as the grounds for physical properties. The panprotopsychist version can be seen as a form of *Russellian neutral monism*, with fundamental protophenomenal properties serving as the grounds for both physical and phenomenal properties. There may also be a mixed view, perhaps *Russellian pluralism*, if some quiddities are phenomenal and some are protophenomenal or unrelated to the phenomenal.¹⁴

On another view, 'mass' refers to a dispositional property that is not grounded in its categorical basis: Instead, categorical and dispositional properties are equally fundamental, and neither is grounded in the other. Given that physical properties are restricted to t-physical properties and those grounded in them, the corresponding version of Russellian monism will be a *Russellian property dualism*, with fundamental physical properties (dispositional properties such as mass) and equally fundamental phenomenal or protophenomenal properties (the corresponding quiddities).

On a final view (the "powerful quality" view advocated by Heil 2012), dispositional properties are identical to their categorical bases. Any corresponding version of Russellian monism will be a version of the Russellian identity theory: Whether 'mass' functions to pick out a dispositional property or its categorical basis, it will pick out a (proto)phenomenal property. One version of this Russellian identity theory (advocated by Mørch 2014) holds that there is a sort of conceptual or a priori connection between (proto)phenomenal properties and the associated dispositions, in the same way that there is arguably such a connection between pain and certain associated dispositions (arguably, one cannot conceive of pain that does not play a certain dispositional role). Another version, which stands to the first version roughly as type-B materialism stands to type-A materialism, holds that there is an a posteriori connection between (proto)phenomenal and dispositional properties. Note that these versions of the Russellian identity theory are consistent with the version discussed a few paragraphs above, on which (for example) 'mass' is equivalent to 'whatever plays the mass role.' They do not entail it, however, as they are also consistent with views on which 'mass' picks out a disposition directly, and they are not entailed by it, as the original version is consistent with views on which dispositional and categorical properties are distinct. One could also see these views as versions of Russellian idealism or neutral monism, on which all truths are grounded in (proto)phenomenal truths.

A number of these versions of Russellian monism differ only verbally. Many of these differences turn on the correct semantics for 'mass' and for 'physical property,' with the underlying metaphysical picture looking the same. One exception here is the difference between Russellian idealism, neutral monism, and pluralism: This turns on the (presumably substantive) issue of whether there is something it is like to have a quiddity. Another may be the differences involving Russellian property dualism and the versions of the Russellian identity theory in the previous paragraph: These turn on the (arguably substantive) issue of whether dispositional properties are grounded in, identical to, or independent of their categorical bases. For what it is worth, I am most attracted to the first version of the Russellian identity theory, with some sympathy also for the idealist, neutral monist, and property dualist versions. The only view that I am entirely unsympathetic with is the a posteriori version of the Russellian identity theory in the previous paragraph (which I think requires a sort of brute identity claim, and so stands to the first version as type-B versions of the familiar identity theory stand to type-A versions). In what follows, I will simply talk of Russellian monism, distinguishing panpsychist and panproto-psychist views as necessary.¹⁵

1.6 Antithesis: The Combination Problem

Given Russellian monism as our new synthesis, a more significant antithesis now threatens. This antithesis takes the form of a major problem for both panpsychism and panprotopsychism: the combination problem.

The combination problem for panpsychism was posed by William James (1890) and named by William Seager (1995). This problem can be stated as follows: How do microexperiences combine to yield macroexperiences? It is at least very hard to see how a number of separate experiences had by separate entities could combine to yield a distinct experience had by a composite entity. It is especially hard to see how they could combine to yield the distinctive kind of macroexperience that we find in our own case.

One way to pose the combination problem is in the form of a conceivability argument. (An approach along these lines is presented by Goff [2009], to whom my argument here is indebted.) Here PP is the conjunction of all microphysical and microphenomenal truths about the world, and Q is a macrophenomenal truth, such as 'Some macroscopic entity is conscious.'

- (1) $PP\&\sim Q$ is conceivable.
- (2) If $PP\&\sim Q$ is conceivable, it is possible.
- (3) If PP&~Q is metaphysically possible, constitutive panpsychism is false.
- (4) Constitutive panpsychism is false.

Here premises (2) and (3) parallel the corresponding premises of the conceivability argument against materialism and are supported by the same reasons. So the key premise here is premise (1). This premise asserts the conceivability of *panpsychist zombies*: beings that are physically and microphenomenally identical to us (and indeed whole worlds that are physically and microphenomenally identical to ours), without any macrophenomenal states.

Why believe that panpsychist zombies are conceivable? Some might find this simply intuitive: One can conceive of all the microexperiences one likes without any macroexperiences. But one can also justify it by invoking a principle in the spirit of James's objection to panpsychism in *The Principles of Psychology*. This is the principle that no set of conscious subjects necessitates the existence of a further conscious subject. Or in the key of conceivability: Given any set of

conscious subjects and any conscious subject not in that set, one can always conceive of all the subjects in the set without the further subject. More precisely: Given any conjunction S of positive phenomenal truths about a group of conscious subjects and any positive phenomenal truth T about a conscious subject not in that group, S&~T is conceivable.

We might say that these principles invoke a subject-subject gap: an epistemic gap from the existence of subjects to the existence of distinct subjects. The principles all have intuitive appeal. Prima facie, it seems conceivable that any group of conscious subjects could exist alone, without any further subjects. But if this is right, constitutive panpsychism is in trouble. Given that all experiences are had by conscious subjects, we can say that microexperiences will be had by microsubjects and macroexperiences by macrosubjects. Then by the principle above, we can conceive of any number of microsubjects having their microexperiences without any macrosubject having macroexperiences. That is, we can conceive of the conjunction of all microphenomenal truths obtaining without any positive macrophenomenal truths obtaining.

This result (along with the conceivability-possibility premise) already rules out a version of constitutive panpsychism on which macroexperience is wholly grounded in microexperience. To rule out all versions, including those in which macroexperience is grounded in microexperience plus physical structure, we can appeal to a modified principle according to which in the case above, S&S'&~T is conceivable, where S' characterizes the physical and structural properties of the members of the original group. This principle seems just about as intuitively plausible as the original principle. Given this principle, premise (1) above follows, and if premises (2) and (3) are granted, constitutive panpsychism is ruled out.

One might think that this problem for panpsychism makes things better for panprotopsychism, as panprotopsychism does not need subjects at the bottom level. Nevertheless, there is also a combination problem for panprotopsychism. This is the problem of how protoexperiences can combine to yield experiences.

As with the combination problem for panpsychism, the combination problem for panprotopsychism can be posed in the form of a conceivability argument. Here *PPP* is the conjunction of all microphysical and protophenomenal truths (or better, purportedly protophenomenal truths, as the combination problem can be used to question whether purportedly protophenomenal properties are truly protophenomenal), and *Q* is a macrophenomenal truth, such as 'Some macroscopic entity is conscious.'

- (1) $PPP\&\sim Q$ is conceivable.
- (2) If *PPP*&~*Q* is conceivable, it is possible.

- (3) If *PPP*&~*Q* is metaphysically possible, constitutive panprotopsychism is false.
- (4) Constitutive panprotopsychism is false.

Once again, the key premise is premise (1). This asserts the existence of *pro-tophenomenal zombies*: beings that share our (purportedly) protophenomenal properties at the microphysical level but that lack consciousness. The conceivability of protophenomenal zombies is perhaps somewhat less obvious than the conceivability of panpsychist zombies, as we have a less clear idea of what protophenomenal properties involve. Still, one might appeal to a general non-phenomenal/phenomenal gap, as on a view I discussed in section 1.4. One thought here is that for any nonphenomenal truths, we can conceive of all these truths obtaining without any experience at all.

Why accept this? One possible justification is a nonsubject/subject gap. This is the claim that no set of truths about nonsubjects of consciousness can necessitate the existence of distinct subjects of consciousness. Or in the key of conceivability: For any set of nonsubjects instantiating nonphenomenal properties and any independent subject exhibiting phenomenal properties, we can conceive of the former without the latter. This principle leads naturally to premise (1) above.

Why believe this principle? One potential justification is the idea that subjects are conceptually fundamental entities. On a view where subjects are metaphysically fundamental entities, then they are not grounded in more fundamental entities, and one can make a case that they are not necessitated by the existence of other fundamental entities. Likewise, if they are conceptually fundamental entities, they are not conceptually grounded in more fundamental entities, and one can make a case that their existence is not a priori entailed by that of other entities. Certainly these principles are not obvious, but they have some intuitive appeal.

Another potential justification is a nonquality/quality gap. Here the idea is that phenomenal properties are qualitative, in that they constitutively involve qualities such as redness, greenness, and so on. And one can argue that nonqualitative truths never necessitate qualitative truths, insofar as one can always conceive of the former obtaining without the latter obtaining. Insofar as purportedly protophenomenal properties are nonqualitative, this principle yields a gap between these properties and the phenomenal that might justify premise (1).

Both panpsychism and panprotopsychism face challenging combination problems, then. As well as sharing a number of problems, each view faces one especially difficult problem that the other does not: the subject-subject gap for panpsychism, and the nonphenomenal-phenomenal gap for panprotopsychism. Reasonable people can differ on which problem is more serious. I am inclined to think the subject-subject problem is more difficult, and that panprotopsychism may benefit from having fewer constraints on its building blocks, but I am far from certain about this. All these problems have the status of challenges rather than refutations, but they are challenges that need to be addressed.

Of course physicalism is faced with its own version of the combination problem: How do microphysical entities and properties come together to yield subjects, qualities, and so on? This challenge is presumably at least as hard as the challenge to panpsychism, as the resources available to the physicalist are a subset of those available to the panpsychist. But we should be clear on the dialectic. The sympathizer with panpsychism has typically already rejected physicalism (at least in non-Russellian forms), precisely on the grounds of these gaps between the physical and the experiential. The question is then whether panpsychism can do any better. It promises to do better in at least one respect: it accommodates the very existence of experience, if only by taking it as fundamental. But it is not clear whether it does any better at explaining the complex manifest character of macroexperience. This is the challenge posed by the combination problem.

By contrast, dualism does not suffer nearly as badly from a combination problem. This is especially clear for substance dualism, which postulates fundamental entities (subjects of experience) that bear macrophenomenal properties. There is no analog of the subject combination problem for such a view. If the dualist takes macrophenomenal properties as fundamental properties, with their structure, qualities, and other features built in, then there will be no analog of the other combination problems either.

Instead of the combination problem, dualism has the familiar problem of mental causation, as well as a problem of economy (why postulate so many fundamental entities?). Panpsychism and panprotopsychism, at least in their constitutive Russellian varieties, do not suffer from these problems. They postulate only as many fundamental entities and properties as are needed to make sense of physics (at least if one thinks that physics requires quiddities), and they make a specific hypothesis about the nature of these properties. And on this picture, phenomenal properties are integrated into the causal order.

I think that substance dualism (in its epiphenomenalist and interactionist forms) and Russellian monism (in its panpsychist and panprotopsychist forms) are the two serious contenders in the metaphysics of consciousness, at least once one has given up on standard physicalism. (I divide my own credence fairly equally between them.) So in a way, our new dialectical situation confronts Russellian monism with (once again) substance dualism. In effect the problems of economy and mental causation for one are weighed against the combination problem for the other. If one of these problems can be solved or proved unsolvable, that will constitute significant progress on the mind-body problem.

1.7 New Synthesis: Panqualityism?

Is a new synthesis in sight? I do not have a solution to the combination problem, so I do not really have a new synthesis. But in this section I want to at least canvas options and to explore one possible new solution, before concluding that it fails. I explore options for dealing with the various aspects of the problem in much more depth in "The Combination Problem for Panpsychism" (Chalmers this volume).

One reaction to the combination problem is to give up on constitutive panpsychism (or panprotopsychism), and instead opt for emergent panpsychism. This view does not face nearly such a pressing form of the combination problem, as it denies that macroexperience is grounded in microexperience. Still, emergent panpsychism loses many of the key advantages of constitutive panpsychism in avoiding the Hegelian dilemma. In particular, it faces a problem of mental causation—how can macroexperience play a causal role?—that is analogous to the problems of dualism and seems to require epiphenomenalism, interactionism, or overdetermination. So it is worth looking closely at the options for constitutive panpsychism.

A second reaction is to hold that macrosubjects are identical to certain microsubjects: that is, they are identical to certain fundamental physical entities with fundamental phenomenal properties, and they share those phenomenal properties. This view avoids the needs for subjects to combine into distinct subjects. One version of this view is akin to Leibniz's 'dominant monad' view, on which human subjects are identical to single fundamental particles, perhaps in their brain. This view is subject to obvious objections, however (What happens if that particle is destroyed? How could a particle have such complex phenomenal properties, especially on a Russellian view?). Another version of the view appeals to fundamental physical entities above the level of the particle: perhaps entangled quantum systems, or perhaps the entire universe. I think that these possibilities (especially the quantum version) are worth exploring, but it is not easy to see how such entities could have fundamental phenomenal properties that yield a phenomenology like ours.

A third reaction is to deflate the subject, either denying that experiences must have subjects at all, or at least denying that subjects are metaphysically and conceptually simple entities. I think it is a conceptual truth that experiences

have subjects: Phenomenal properties must be instantiated by something, and they characterize what it is like to be that thing. But the second denial seems more tenable. Indeed, some such denial seems required to be a constitutive panpsychist, a constitutive panprotopsychist, or indeed a materialist. This view may require rejecting certain intuitions about subjects, but these intuitions are not nonnegotiable.

We might define Subjects as primitive subjects of experience. I think that we have a natural conception of Subjects: These are subjects as they might have been in the Garden of Eden, as it were. Where Subjects are concerned, the subject-subject gap and the nonsubject-subject gap are both extremely plausible: The existence of a Subject is not necessitated or a priori entailed by the existence of distinct Subjects or indeed by the existence of non-Subjects. So if we are Subjects (and if we set aside the view that macrosubjects are identical to microsubjects), constitutive panpsychism and constitutive panprotopsychism are false.

Still, it is far from obvious that we are Subjects. There does not seem to be an introspective datum that we are Subjects, and it is not obvious that there are strong theoretical arguments to that effect. There are perhaps intuitions of determinacy about personal identity that tend to support the claim (see Barnett 2010; Nida-Rümelin 2010), but these intuitions do not seem to be nonnegotiable, and there are reasonably strong considerations in favor of rejecting them (see Parfit 1984). And once we deny that we are Subjects, the door is at least opened to rejecting the subject-subject gap and the nonsubject-subject gap, and to accepting constitutive panpsychism or panprotopsychism.

I think that a Russellian monist must almost certainly embrace this view (perhaps the only remotely promising alternative is the quantum version of the micro/macro identity claim above). Still, to deny that we are Subjects is not to solve the combination problem. We still need to give an account of how macroexperience can be grounded in microexperience or in protoexperience.

Here I will look briefly at a view that has been popular among sympathizers with panpsychism and panprotopsychism: panqualityism. The name of this view was introduced in an article by Herbert Feigl (Feigl 1960), who credits the term to conversation with Stephen C. Pepper, but versions of the view itself were popular among the neutral monists of the early twentieth century, including William James, Ernst Mach, and Bertrand Russell (James 1904; Mach 1886/1959; Russell 1921). More recently, the view has been defended by Sam Coleman (2012).

On this view, *qualities* are the properties presented in experience: Intuitively, these are properties like redness, greenness, heat, and so on. Qualities are not identical with phenomenal properties: When redness is presented to me in experience, I have a phenomenal property, but I need not be red. Instead, we

would intuitively say that I am aware of redness, and that phenomenal properties involve awareness of qualitative properties. Likewise, phenomenal properties are always instantiated by conscious subjects, but qualities need not be. We can certainly make sense of the idea of a red object that is not a subject of experience.

Panqualityism typically requires rejecting a reductionist view of qualities, such as a view on which color qualities are identified with physical reflectance properties or something of the sort. Instead, it is naturally associated with what I have called an Edenic view of qualities. Here the qualities most fundamentally presented in experience are properties such as Edenic redness, a simple property that may not be instantiated by the objects that seem to have it in the external world, but which might have been instantiated in the Garden of Eden.

Panqualityism holds that fundamental physical entities instantiate qualities like these. We might imagine, for example, that fundamental particles are edenically red. More likely, the relevant qualities involved will be more austere than this, but they will nevertheless be primitive properties that could be presented in experience. The most important kind of panqualityism, unsurprisingly, is constitutive Russellian panqualityism, on which qualities serve as quiddities and also serve to constitute human experience. Many of the panqualityists discussed above have endorsed views of this sort.

Constitutive panqualityism is a form of panprotopsychism rather than panpsychism: Qualities are not phenomenal properties but serve to constitute phenomenal properties. Because qualities need not be instantiated by subjects, the view need not invoke microsubjects at all. Panqualityism is occasionally characterized as a version of panpsychism with 'experiences without subjects' or 'unsensed sensa,' but I think the view is best regarded as a form of panprotopsychism. Still, it is a view on which the protophenomenal properties take an especially familiar form, and on which they have a close connection to phenomenal properties.

Panqualityism is not threatened by the subject-subject gap, as it does not require microsubjects to constitute macrosubjects. Likewise, it is not threatened by the nonquality-quality gap, as the purportedly protophenomenal properties here are qualitative through and through. It is threatened by the nonsubject-subject gap, but here it responds by deflating the subject. Some of the traditional panqualityists rejected subjects of experience altogether, while others have taken deflationary views of them on which they can be constituted by underlying qualities, perhaps along with structural properties.

How does panqualityism solve the combination problem? It is natural for the panqualityist to argue that simple microqualities can collectively constitute complex macroqualities, ultimately building up something as complex as the qualitative structure of a visual field or even a full multisensory field. Then it could be suggested that the existence of these complex qualities explains the phenomenal data even without postulating an associated subject of experience; or it could be suggested that certain complex qualities entail the existence of an associated subject, perhaps in a deflated sense.

Still, I think that panqualityism is vulnerable to a version of the combination problem analogous to earlier versions. In particular, we can mount a conceivability argument against panqualityism as follows. Here QQ is a conjunction of positive qualitative truths at the microphysical, perhaps along with any other microphysical truths, and Q is a positive macrophenomenal truth.

- (1) $QQ\&\sim Q$ is conceivable.
- (2) If QQ&~Q is conceivable, it is metaphysically possible.
- (3) If QQ&~Q is metaphysically possible, constitutive panqualityism is false.

(4) Constitutive panqualityism is false.

Again, all the action is in the first premise. This premise asserts the conceivability of *qualitative zombies*, beings that are qualitatively (and microphysically) identical to us without consciousness.

Why believe this premise? One could make a case that it is intuitively obvious. But more deeply, it is grounded in what we might call the qualityawareness gap. Here the idea is that no instantiations of qualities ever necessitate awareness of qualities. Or in the key of conceivability: For any set of instantiated qualities and physical properties, it is conceivable that all those qualities and properties are instantiated without any awareness of the qualities. Given that all phenomenal properties involve awareness of qualities, premise (1) above follows. And even if only some phenomenal properties involve awareness of qualities, this will be enough to make the case against constitutive panqualityism.

The quality-awareness gap has much intuitive force. On the face of it, it is conceivable that Edenic redness is instantiated without anyone being aware of it. And on the face of it, this intuition scales up to arbitrarily complex qualities. Even given complex qualities corresponding to the structure of a visual field, then if it is conceivable that those qualities be instantiated at all (presumably by a situation in the world corresponding to the situation as perceived), it is conceivable that they be instantiated without any awareness of those qualities.

The panqualityist might respond in various ways. They could bring in awareness at the fundamental level, perhaps by appealing to special qualities that cannot be instantiated without awareness of those qualities (pain, maybe?); but this leads back to subjects at the fundamental level and the associated problems. They might deny the existence of awareness, as James (1904) does, and hold that our experience involves qualities but does not involve awareness of them; but this claim runs directly counter to our phenomenology. They might combine the appeal to qualities with a functional reduction of awareness, as Coleman (2012) does; but I think that the conceivability argument above itself gives reason to reject such a reduction.

Panqualityism is also vulnerable to other aspects of the combination problem. It is vulnerable to the *structure combination problem*: The structure among qualities instantiated in the brain is very different from the structure among qualities of which we are aware, and it is hard to see how the former could constitute the latter. It is also vulnerable to the *quality combination problem*: It is hard to see how a few primitive qualities (which is all that the Russellian panqualityist can appeal to) could yield the vast array of qualities of which we are aware.

I conclude that panqualityism does not offer a solution to the combination problem. We are still in need of a new synthesis.

1.8 Conclusion

We started with the thesis of materialism and the antithesis of dualism, and reached the synthesis of panpsychism. This synthesis encountered the antithesis of panprotopsychism, from which we reached the new synthesis of Russellian monism. This synthesis encountered the antithesis of the combination problem, and whether there can be a new synthesis remains an open question.

Still, I think that the Hegelian argument gives good reason to take both panpsychism and panprotopsychism very seriously. If we can find a reasonable solution to the combination problem for either, this view would immediately become the most promising solution to the mind-body problem. So the combination problem deserves serious and sustained attention.

Notes

- I first presented this material at the Munich conference on panpsychism and emergence in June 2011. I am grateful to the audience there and also to audiences at Amherst, Bogazici, Charleston, Fordham, Notre Dame, Santiago, Stanford, and Wesleyan. I owe a special debt to Daniel Stoljar whose related work in "Two Conceptions of the Physical" greatly influenced this chapter. Thanks also to Torin Alter, Sam Coleman, Brian Garrett, Philip Goff, John Gregg, Bill Meacham, Daniel Stoljar, Galen Strawson, and Keith Turausky for comments on this chapter.
- 2. I gather that in fact this dialectical form comes from Fichte, and that Hegel dismissed it as simplistic. Still, I will stay with the popular attribution.
- 3. For more on the notion of grounding, see Schaffer (2009) and Fine (2012). The notion of grounding at play here is what is sometimes called 'full grounding,' involving a 'wholly in

virtue of' relation, as opposed to 'partial grounding,' which involves a 'partly in virtue of' relation. The latter is inappropriate for defining materialism, as the definition would then allow a nonmaterialist view on which truths about consciousness obtain in virtue of physical truths along with some other nonphysical truths.

- 4. For a much-elaborated version of the argument using two-dimensional semantics, see Chalmers (2010).
- 5. Principles such as (3) are sometimes put with 'is a physical property' instead of 'is grounded in a physical property.' This amounts to an overly strong causal exclusion claim on which high-level events and their low-level grounds cannot both be causally relevant. Reflection on standard cases (Bennett 2003; Yablo 1992) suggests that constitutively connected events need not exclude each other as causes: These are cases of 'benign overdetermination' as opposed to cases of 'bad overdetermination.' Premise (3) excludes only cases of the latter sort.
- 6. The distinction between narrowly and broadly physical properties is closely related to Stoljar's distinction between t-physical properties (properties invoked by physical theory) and o-physical properties (intrinsic properties of physical objects), but it is not the same distinction. For a start, given a view on which 'mass' refers to a quiddity that plays the mass role, mass will be t-physical (assuming a property is invoked by physical theory iff it is referred to by an expression of that theory) but it will not be narrowly physical. And given a view on which physical objects have epiphenomenal intrinsic properties that are not those invoked by physical theories and that are not quiddities, these properties will be o-physical but will not be broadly physical. For related reasons (discussed later), I think the broad/narrow distinction is better suited than the t-/o- distinction to do the work that Stoljar wants the latter to do. Note that in Chalmers (2010, 192) I use the broad/narrow terminology to mark a different distinction.
- 7. Structural and categorical zombies are closely related to the t-zombies (t-physical duplicates without consciousness) and o-zombies (o-physical duplicates without consciousness) discussed by Stoljar (2001a). As before I think the broad/narrow distinction is more crucial than the t-/o- distinction here.
- 8. Alternatively, the constitutive Russellian panpsychist can accept premise (2) asserting the causal closure of the narrowly physical, while rejecting premise (3). They can hold that the narrowly physical explanation is itself grounded in a broadly physical explanation, so that these explanations are not independent and a bar on overdetermination does not render them incompatible. The case for premise (3) tacitly assumes that physical explanations do not themselves have further grounds; but on a Russellian view, narrowly physical explanations have further grounds.
- 9. The Hegelian argument could in principle be formalized as a six-premise argument that uses the three premises of the conceivability argument (with 'physical' disambiguated to mean narrowly physical) and the three premises of the causal argument (with 'physical' to mean broadly physical) to establish the conjunction of broad physicalism with the denial of narrow physicalism. An argument structure along these lines is at play in Stoljar (2001b, section 4), with the main differences being that Stoljar invokes the knowledge argument rather than the conceivability argument, uses the o-/t- distinction where I use the broad/ narrow distinction, and rejects panpsychism.
- 10. What about type-B views that appeal to quiddities that satisfy (i) but not (ii)? Some such views may nevertheless have a 'panprotopsychist' flavor, perhaps because of the special flavor of the quiddities they appeal to, while others (say, the view advocated by Papineau 2002) seem to lack that flavor. A line between these views is hard to draw, so for present purposes I count none of them as panprotopsychism.
- 11. Stoljar (2013) credits the first appearance of "Russellian Monism" to Chalmers (1999). The phrase actually appears (alongside "panprotopsychism") in Chalmers (1997), but in any case it is heavily inspired by Michael Lockwood's talk of the "Russellian identity theory" in *Mind*, *Brain*, *and the Quantum* (Lockwood 1989).
- 12. Stoljar and Strawson are naturally counted as expansionary Russellian physicalists. Strawson spends some time arguing with people like me (e.g., questioning whether physical duplicates without consciousness are conceivable), but once it is clear that I mean by

'physical' what Strawson means by 'physics-al,' the disagreement between us may become largely verbal.

- 13. Maxwell (1978) and Lockwood (1989) are certainly Russellian identity theorists: Both explicitly endorse the identity theory and credit the underlying idea to Russell. Feigl (1958) and Montero (2010) can easily be interpreted as holding the view. The coherence of the Russellian identity theory, on which quiddities are identical to t-physical properties and on which t-physicalism is true, suggests that Russellian monism is not best characterized (following Stoljar) as o-physicalism about consciousness without t-physicalism.
- 14. Bolender (2001) puts forward a sort of Russellian idealism, and Rosenberg (2004) may be either a Russellian idealist or pluralist.
- 15. The different versions of Russellian monism will take different attitudes to the conceivability and possibility of zombies: physical duplicates without phenomenal states. Expansionary Russellian physicalism will deny that they are conceivable or possible: Given the expanded sense of the physical, to conceive of a zombie requires conceiving of a categorical zombie (same structure, same quiddities, no consciousness), which cannot be done according to the view. Russellian identity theorists of the first sort discussed above may hold that zombies are conceivable but not possible, because of nontrivial two-dimensional structure in physical terms (the primary intension of 'mass' picks out whatever plays the mass role, the secondary intension picks out the quiddity that actually plays the mass role). Russellian idealists, neutral monists, and property dualists may well hold that zombies are conceivable and possible, in that there are conceivable and possible situations where the structural properties are associated with different quiddities that are independent of the phenomenal or perhaps with no quiddities at all.

Bibliography

- Barnett, David. "You Are Simple." *The Waning of Materialism*. Ed. Robert Koons and George Bealer. New York: Oxford University Press, 2010.
- Bennett, Karen. "Why the Exclusion Problem Seems Intractable and How, Just Maybe, to Tract It." *Noûs* 37.3 (2003): 471–97.
- Bolender, John. "An Argument for Idealism." Journal of Consciousness Studies 8.4 (2001): 37-61.
- Chalmers, David. "Moving Forward on the Problem of Consciousness." *Journal of Consciousness Studies* 4.1 (1997): 3–46.
- Chalmers, David. "Précis of 'The Conscious Mind. Philosophy and Phenomenological Research." *David Chalmers*. 1999. Web.
- Chalmers, David. "Consciousness and Its Place in Nature." *Blackwell Guide to the Philosophy of Mind*. Ed. Stephen Stich and Ted Warfield. Oxford: Wiley-Blackwell, 2003.
- Chalmers, David. "The Two-Dimensional Argument Against Materialism." *The Character of Consciousness.* Ed. David Chalmers. New York: Oxford University Press, 2010. 141–91.
- Chalmers, David. Constructing the World. New York: Oxford University Press, 2012.
- Chalmers, David. "The Combination Problem for Panpsychism. *Panpsychism*. Ed. Godehard Brüntrup and Ludwig Jaskolla. New York: Oxford University Press, this volume.
- Coleman, Sam. "Mental Chemistry: Combination for Panpsychists." *Dialectica* 66.1 (2012): 137–66.
- Feigl, Herbert. "The 'Mental' and the 'Physical." Minnesota Studies in the Philosophy of Science. Vol 2. Ed. Herbert Feigl, Michael Scriven, and Grover Maxwell. Minnesota: University of Minnesota Press, 1958. 370–497.
- Feigl, Herbert. "Mind-Body, Not a Pseudo-Problem." *Dimensions of Mind*. Ed. Sidney Hook. New York: Collier Books, 1960. 24–36.
- Fine, Kit. "The Pure Logic of Ground." *Review of Symbolic Logic* 5.1 (2012): 1–25.
- Goff, Philip. "Why Panpsychism Doesn't Help to Explain Consciousness." *Dialectica* 63.3 (2009): 289–311.

Heil, John. The Universe as We Find It. New York: Oxford University Press, 2012.

- James, William. The Principles of Psychology. Boston: Henry Holt, 1890.
- James, William. "Does 'Consciousness' Exist?" Journal of Philosophy, Psychology, and Scientific Methods 1.18 (1904): 477–91.
- Lockwood, Michael. Mind, Brain, and the Quantum. New York: Oxford University Press, 1989.
- Mach, Ernst. 1886. *The Analysis of Sensations and the Relation of Physical to the Psychical*. Trans. C. M. Williams. Mineola, NY: Dover, 1886/1959.
- Maxwell, Grover. "Rigid Designators and Mind-Brain Identity." Minnesota Studies in the Philosophy of Science 9.9 (1978): 365–403.
- Montero, Barbara. "A Russellian Response to the Structural Argument Against Physicalism." Journal of Consciousness Studies 17.3-4 (2010): 70-83.
- Nida-Rümelin, Martine. "An Argument from Transtemporal Identity for Subject-Body Dualism." *The Waning of Materialism.* Ed. Robert Koons and George Bealer. New York: Oxford University Press, 2010. 191–214.
- Papineau, David. Thinking About Consciousness. New York: Oxford University Press, 2002.
- Parfit, Derek. Reasons and Persons. New York: Oxford University Press, 1984.
- Rosenberg, Gregg. A Place for Consciousness: Probing the Deep Structure of the Natural World. New York: Oxford University Press, 2004.
- Russell, Bertrand. The Analysis of Mind. London: George Allen and Unwin, 1921.
- Russell, Bertrand. The Analysis of Matter. London: Kegan Paul, Trench, Trubner, 1927.
- Schaffer, Jonathan. "On What Grounds What." Metametaphysics: New Essays on the Foundations of Ontology. Ed. David Chalmers, David Manley, and Ryan Wasserman. New York: Oxford University Press, 2009. 347–83.
- Smart, Jack. "Sensations and Brain Processes." Philosophical Review 68.2 (1959): 141-56.
- Seager. William. "Consciousness, Information, and Panpsychism." Journal of Consciousness Studies 2.3 (1995): 272–88.
- Stoljar, Daniel. "The Conceivability Argument and Two Conceptions of the Physical." Philosophical Perspectives 15 (2001a): 393-413.
- Stoljar, Daniel. "Two Conceptions of the Physical." *Philosophy and Phenomenological Research* 62.2 (2001b): 253-81.
- Stoljar, Daniel. "Four Kinds of Russellian Monism." *Current Controversies in the Philosophy of Mind*. Ed. Uriah Kriegel. New York: Routledge, 2013.
- Strawson, Galen. "Realistic Monism: Why Physicalism Entails Panpsychism." Journal of Consciousness Studies 13.10–11 (2006): 3–31.
- Yablo, Stephen. "Mental Causation." Philosophical Review 101.2 (1992): 245-80.

2

Emergent Panpsychism

GODEHARD BRÜNTRUP

2.1 Outline

"Se non è vero, è molto ben trovato—If it is not true, it is well conceived." This quote is usually attributed to Giordano Bruno (1548–1600). Bruno was one of the great panpsychists of the Renaissance. He believed that all matter is permeated by spirit. Panpsychism has always been a player in Western metaphysics, albeit not as prominent as dualism or materialism. The mind-body problem might be inscrutable to us and panpsychism might not be true, but it is a coherent position that deserves more attention in the contemporary debate. In this chapter three claims will be defended:

- 1. Panpsychism is a genuine and conceptually clearly delineated position in the philosophy of mind. It should be taken seriously as an alternative to the dominant and somewhat simplistic view that the only viable options are physicalism or dualism. It is distinct from idealism and neutral monism as well.
- 2. Constitutive panpsychism is currently the preferred variant of panpsychism. Retaining the causal closure of the physical, this compositional form of panpsychism seems to provide a metaphysics of mental causation that is acceptable within a broadly physicalist framework. It is doubtful, however, whether it can preserve a robust metaphysical account of agency. This is the reason why alternative versions of panpsychism deserve attention.
- 3. Nonconstitutive panpsychism, a hitherto less developed account, has the resources to overcome some of the problems facing constitutive panpsychism. Even if incorporating the notion of downward causation, it is nevertheless clearly distinct from substance dualism.

2.2 Panpsychism as a Distinct Position in the Philosophy of Mind

This section will sketch the most general metaphysical claims of panpsychists and attempt to differentiate panpsychism from the more general intuitions of Russellian monism.

Surprisingly, panpsychism is often accused of being physicalism or dualism in disguise, that is, of collapsing into one of these two positions in the final analysis. The surprising fact that both of these claims are actually advanced can be explained more easily by recalling that two distinct versions of panpsychism are distinguished. One, constitutive panpsychism, is closer to classical physicalism; the other, nonconstitutive panpsychism, is closer to classcal dualism. Both versions of panpsychism share basic assumptions that set them apart from these other accounts. Constitutive panpsychism claims that macroexperience is constituted by, grounded in, or realized by microexperience. Nonconstitutive panpsychism claims that unified macroexperience is an emergent phenomenon that cannot be fully captured by the metaphysical concepts of constitution, composition, grounding, or realization.

2.2.1 Ontological Simplicity

Like other metaphysicians, panpsychists are often motivated by certain general and overarching metaphysical ideals of simplicity and austerity: the homogeneity of nature, monism, and a single-category ontology. Many physicalists and idealists share the same goal of theoretical simplicity, including proponents of a physicalist event ontology or an idealist Leibnizian monadology. From the point of the panpsychist, physicalism and idealism do, however, pay a substantial metaphysical price by either reducing the mental to the physical or by deflating the physical to the status of a well-founded phenomenon. Taking both the mental and the physical seriously, while avoiding the complexities of substance dualism, is certainly a major motivation for panpsychists. If panpsychism cannot preserve the elegant simplicity of physicalism or idealism, then dualism might well be the more attractive alternative, since it is capable of preserving our common-sense worldview more effectively than panpsychism.

The homogeneity thesis claims that nature is built up from the same kind of basic entities through space, and even through time. Beings in very distant galaxies will be made up of basic elements that are prevalent in our part of the universe. So it will not be the case that some cosmic locations will contain simple Cartesian souls, whereas others contain only elementary physical particles and compositions of particles. According to the panpsychist, it is also not the case that some cosmic locations contain entities endowed with experience and that the entities in others are completely void of even the simplest form of experience. Some experiential aspect (which might be very primitive and not necessarily conscious) is fundamental to, and ubiquitous in, concrete particulars. Experience (in the widest possible sense) exists not only at some rare level of complexity, such as animal brains, but is in some form present at all levels of nature. The diachronic ideal of homogeneity states the same idea in temporal terms. In the history of the universe there is no clear cut-off point at which experience emerges out of a past that is absolutely void of any experience. Rather, experience was somehow "present at the very origin of things" (James 1983, 152); it is a fundamental feature of nature.

The theoretical intuition guiding these thoughts is the so-called genetic argument. Ex nihilo nihil fit, nothing can emerge out of nothing. More precisely, it states that nothing can give something which it does not possess. For something to emerge out of something else, the former must have been present in some diminished form right from the beginning. For example: Even if the solidity and impenetrability of macroobjects cannot be found in the microworld, there have to be physical properties that serve as an emergence base for these macroproperties; otherwise the emergence becomes a mere brute and unintelligible fact. Imagine a Platonic world inhabited only by abstract entities. The idea that in this world some kind of configuration of abstract objects can cause the emergence of a concrete material object seems unintelligible. If there are no spatio-temporal entities in a given world from the outset, no configuration or structure in that world will bring about spatio-temporal particulars. Similarly, the panpsychist claims that the emergence of phenomenal minds out of a world which lacks a scintilla, even the faintest form, of experience, is simply unintelligible. One might also call this the denial of inter-attribute emergence. Say, a given ontology contains these four basic attributes: concrete, abstract, physical, mental. If, for example, a given possible world in this ontology contains at its foundation concrete physical objects only, then neither mental nor abstract entities can emerge in it, at least it cannot be made intelligible how they could emerge. But higher levels of concrete physical objects can emerge in it because this would only amount to a case of intra-attribute emergence.

2.2.2 Monism

For the same reason, panpsychism is often labeled as a kind of monism. But this description might be too coarse-grained. Of course, seen as an alternative to Cartesian dualism, panpsychism is monistic in spirit. There is a possible reading of panpsychism where all physical properties are grounded in mental

properties. If grounding means that the intrinsic properties are the constitution base of the relational extrinsic properties, then we have a metaphysical system in which certain fundamental entities with absolutely intrinsic properties constitute all the remainder of reality-a view that Leibniz famously argued for in his Monadology. The whole of physical space was constructed by giving each monad a spatial viewpoint from which space was constituted as experienced space. But then panpsychism simply collapses into idealistic monism. The version of monism that comes closest to panpsychism is neutral monism, but neutral monists take great care to distinguish their position from panpsychism. Neutral monism is often somewhat Kantian in spirit, claiming that we do not know the ultimate nature of the universe but only the mental and the physical arising from that ultimate nature. The neutral is simply not known to us; it is a postulate of reason. This Kantian ignorance about the ultimate nature of things was certainly a dominant motive for Russell's neutral monism. While Russell flirted with the idea that the intrinsic properties of the physical might be mental in nature, his predominant attitude about the intrinsic nature of things seems to have been agnosticism. If panpsychism is labeled as 'Russellian monism,' then—while not being a complete misnomer—relevant differences will be obscured. The panpsychist would argue that an entirely neutral basis that does not even contain some primitive or 'proto' form of mentality would not help us at all in explaining the emergence of the mind. Thus panpsychistic monism is a 'dual-aspect monism,' claiming that there is only one kind of thing but it features physical and mental properties. Dual-aspect monism is not a contradiction in terms, but the concept certainly makes it clear that panpsychism is neither monism nor dualism simpliciter. Panpsychism is more complicated than these crude alternatives suggest. In the contemporary debate about panpsychism, the mental part is usually constituted by the intrinsic properties of things (the quiddities), whereas the physical part is constituted by the relational-structural properties. Whether this is still a form of monism in the full sense is disputable. One could also, with some justification, call it 'dualism all the way down.' Panpsychism is not a monism tout court as is physicalism or idealism. Even so, it may correctly be labeled 'monist' in the sense that it can be construed as a one-category ontology. Whitehead's panexperientialism is a good example for this. For Whitehead there is only one basic category of entities: events with both a physical and mental pole. He calls events 'actual occasions.' They are the most fundamental entities. Everything else is constructed out of them. The question, however, is whether this construction allows for emergent novelty rather than merely a resultant composition of existing entities. Whitehead was a nonconstitutive panpsychist who allowed for the strong emergence of new individuals. These new individuals were strongly emerging actual entities endowed with a new mental unity and subjectivity.

In the most recent debates, panpsychists have often tried to avoid strong emergence. The attraction of constitutive panpsychism is rooted in its ability to preserve a broadly reductionist picture. The microentities determine all facts. But not all facts are physical facts in the narrow sense. There is a categorical basis that carries the relational properties described by physics; ignoring it is the reductive physicalism's mistake. This move enables the panpsychist to escape the zombie argument against physicalism by claiming that a complete copy of the physical world must include the quiddities and not just the relational properties. Such a metaphysically complete copy of the lowest physical level—that is, a copy which is not only structurally isomorphic but copies the intrinsic natures as well-will give rise to consciousness with the strong necessity that the relation of logical supervenience provides. It might seem that constitutive panpsychism is simply physicalism in disguise. But to think this would be a grave misunderstanding, conflating two senses of 'physical.' To use Strawson's terminology, it is important to conceptually distinguish 'physicalism' from 'physicSalism.' The metaphysical work that is being done by intrinsic mental properties, even at the most basic level of nature clearly distinguishes constitutive panpsychism from physicalism.

2.2.3 The Carrier Thesis

The distinction between physicalism and physicSalism points again to the duality inherent in panpsychism, even if the latter is construed as a one-category ontology. The physical structure is metaphysically incomplete, requiring a categorical base. Only the composition of both the structural and the non-structural will give rise to a real concrete entity. This thesis could be called the 'hylomorphic thesis' or better the 'carrier thesis.' As in Aristotle's metaphysics, relational structure can only exist together with something ultimately non-structural which it configures. But the term 'hylomorphism' is strictly associated with the Aristotelian fundamental duality of form (*morphe*) and matter (*hyle*). In the Aristotelian tradition form requires something which it configures (prime matter). Aristotelian prime matter carries the relational structure of the world. Structure alone is never sufficient for the existence of a concrete particular. Panpsychists often argue in a parallel way, but for them the carrier is not prime matter but something mental or analogous to the mental.

The modern post-Cartesian view of matter reduces matter to something that can be fully described by mathematical relations in algebra, analytic geometry, and calculus. It thus focuses on the structural properties of things. Even the Cartesian idea of matter as mere extension is silent on the question of what is being extended. Extension is merely the abstract notion of the repetition of something—a point that was already made by Leibniz (see Brüntrup 2009, 246). Leibniz argues that the concept of extension cannot help explicate the nature of the substance that is being spread out, and that, on the contrary, substance is ontologically prior to the repetitive multiplicity of extension (see Leibniz G IV, 467). Formally speaking, extension is the repetitive multiplicity of point-like entities. But the nature of those entities is specified in the sciences simply by the relations into which they enter.

A pragmatic way to avoid these deep metaphysical waters would be to resort to epistemic structuralism, which is antirealist in spirit and remains agnostic about any nonstructural properties of the unobservable. Nature, as we know by empirical investigation, is only nature as it presents itself through the mathematical analysis of sense data. The Russell of "The Analysis of Matter" provides a good example of this view (see Russell 1927). Russell argued for an agnosticism concerning the physical world, with the exception of its purely formal and mathematical properties: ". . . it would seem that wherever we infer from perceptions it is only structure that we can validly infer; and structure is what can be expressed by mathematical logic" (Russell 1927, 254). And: "The only legitimate attitude about the physical world seems to be one of complete agnosticism as regards all but its mathematical properties" (Russell 1927, 270). Higher order properties of physical theories can only be expressed in mathematical terms. Epistemological structural realism limits the scope of scientific realism to exactly these properties. But according to the epistemic structural realist there is an objective world out there that contains unobservable objects, of which we can only know the relational properties. Thus we only know the structure, ultimately only the formal structure, of the world. This idea, unsurprisingly, is much older than Russell. Kant argued that things in the phenomenal world are wholly constituted by their relations. He considers it a conceptual truth, however, that things as objects of pure understanding must have intrinsic properties. Even in the 'postmetaphysical' "Critique of Pure Reason" Kant argues that those intrinsic properties must be analogous to the ones presented to me by my inner sense. "They must be something which is either itself a thinking or analogous to thinking" (Kant CPR, B321). This is a metaphysical argument that most panpsychists would gladly embrace (see Brüntrup 2011, 24).

The classic critique of epistemological structural realism is "Newman's Argument" (Newman 1928, 139–140), which was initially directed against Russell. Newman's Argument is best understood as a *reductio* aimed at epistemological structural realism, showing that it ultimately collapses into antirealism. The existence of a structure is trivially true of a set of objects. According to Newman, a statement describing a certain structure with regard to a number of objects is trivial. Why is it trivial to claim that a set of objects has a (or some particular) structure? Because, for Newman, a structure is purely formal and

mathematical, and is furthermore independent of the intrinsic qualities of the objects. If only the structure is known, then beside what is logically deducible from the properties of the structure, the only thing that can be known is the number of its constituting objects. But if all we know about the objects is their cardinality, that is, if we do not know any properties of the objects that ground certain relations and exclude others, then—mathematically—any system of relations over these objects is as good as any other; all of them are instantiated. Relations are simply sets of ordered sequences of entities. Given the entities, all of those ordered sequences will exist, as a matter of pure mathematics. If Newman's argument is correct, then not only do we have no knowledge of the intrinsic properties of things, we do not even know the objective structure of the world in any realist sense of 'objective.' Scientific realism understood as epistemic structural realism collapses into antirealism.

The same point can be made of Putnam's famous model-theoretic argument (see Putnam 1980). As Newman argues, given a number of objects, any relational structure configuring them is already given. If we picture objects as mere nodes in a relational graph and as having no intrinsic nature, then for each structure there are many different relations between the objects that make true the propositions describing that structure. What is the intended model of the structure? Which is the one and only relational structure of the mindindependent world? There are too many ontological interpretations (models) for our theories. Our scientific descriptions of the world are unable to single out the intended model, that is, the real world. Since science deals only with mathematical structure and not with the relations which are determined by the qualitative intrinsic natures of the relata, we can never know the one true story about the world in a metaphysical-realist way. We have too many 'truths.' Panpsychists are not well advised to take this broadly Russellian route, which leads not only to agnosticism about the intrinsic properties of concrete entities but to antirealism in the philosophy of science.

2.2.4 Intrinsic Natures

Russell's views do not fit smoothly into the metaphysical debates within the contemporary analytic philosophy of mind. There is a significant element of idealism in Russell's thought which simply does not square easily with the metaphysical realism prevalent in the current debate. The Russell of "Our Knowledge of the External World" (Russell 1914) is still present in his writings from the 1920s. For the early Russell the world does not cause sense data; rather, sense data construct the physical object. It is not far from this move to the Kantian claim that the intrinsic nature of physical reality is unknown to us. Genuine panpsychism, as understood here, is a robust metaphysical thesis

in which scientifically inaccessible intrinsic mental properties (or properties somehow analogous to mental ones) play an important role in grounding the relational structure of the world. The relational properties of substances must have an underlying foundation in intrinsic properties. Famously, Leibniz claimed in his "Letter to de Volder" that there is no denomination so extrinsic that it does not have something intrinsic as its foundation (Leibniz G I, 240). If this is correct, then we need ultimate intrinsic properties that carry the entire existing net of functional-relational properties.

This intuition resurfaces in contemporary debates. John Haugeland endorses the traditional view that a substance needs certain properties which it maintains regardless of anything else. He considers the ontological status of the pieces in a chess game—say a rook or a pawn—and claims that their very nature is determined entirely by how they move about in the game in relation to other pieces. "No rook is a substance. ... Nothing about a rook is determinate, not even its 'rookness,' apart from its participation in a chess game" (Haugeland 1993, 63). The formal definition of a type in a chess game is circular. The nature of each type is completely determined by its set of allowable moves within the game as a whole. The chess game as a whole, however, is defined by the interdependent set of types which play functional roles within it. Each part of the game presupposes the existence of the whole game, and the game presupposes the existence of its parts. Why isn't this circularity of chess categories vicious? How can chess games actually and concretely exist? Classic functionalism has an answer to this question, which is quite similar to the one given by Aristotelian hylomorphism: There must be something distinct from the formal structure that actually grounds that structure in concrete reality. In the case of a chess game we have physically distinct objects that serve as standins or realizers of the relevant types, thus allowing for the existence of concrete tokens of those types.

Of course, there is much more to consider here, like the concrete chessboard or the physical position of the players in space. Without such 'carriers' of the formal structure, the game would remain too incomplete and abstract to exist concretely. The panexperientialist Gregg Rosenberg extends this thought to other, more complex, conceptual systems such as those constructed by scientific theories (see Rosenberg 2004, 234). Such systems too are abstract and circularly defined. Consider cellular automata in computer science. These are like giant chess games. Each cell is defined by its role in the entire system, and the entire system is defined by the cells. Cellular automata may exist as computational systems because there is something external to the formal system that realizes or carries them. The physical states of the hardware are the carriers of the cellular automata as functional systems. Biology as an abstract functional system is carried by the mechanics of molecular biochemistry, psychology by the dynamical properties of the neural system, economics by the needs and desires of individuals. The crucial question, however, is: What carries the most basic physical level? Physics presents us with a world of interdependently defined functional roles. Are there any properties that can give this circularly defined conceptual system a foothold in concrete reality? This is a puzzling question that Rosenberg calls the 'ultimate-carrier problem.' It is very similar to the question Leibniz raised with regard to Descartes's notion of matter.

There are good scientific reasons to assume that nature has a lower size limit (i.e., a Planck-size scale). In order to avoid an infinite regress of ever more finegrained systems, where each lower structure serves as the carrier of the next higher one, a stopper is needed. Only a property that is intrinsic *tout court* and not relative to a system could bring this about and serve as an ultimate carrier. Are there properties that are not intrinsic to any system but at least partly intrinsic to themselves? The only candidates we know of are phenomenal qualities, or something analogous to phenomenal experience. "Analogous" means that these properties might be vastly dissimilar to higher-level phenomenal properties, for similarity is not a transitive relation. Even if mental properties of adjacent layers of nature are similar, the mental properties of nonadjacent layers may be quite dissimilar. But they must have something in common with the phenomenal properties we experience. They cannot be understood in purely relational terms. One cannot understand the nature of these phenomenal qualities by knowledge of their contextual relations alone. Radical intrinsicness is the very nature of phenomenal qualia. Whatever grounds the structural-relational properties of the world must have this radical intrinsicness. It might well be that our own consciousness is the closest analogue we have to this underlying reality. This is the so-called argument from intrinsic natures for panpsychism. As we have seen, even Kant agrees with this line of thinking, provided that what we are seeking is a metaphysical truth. In his critical philosophy, however, he opts for an epistemic constraint within the boundaries of possible experience. A similar skepticism is present in Russell's agnosticism about the ultimate intrinsic nature of concrete entities. The panpsychist, however, cannot enjoy the luxury of withholding judgment here. For the panpsychist, at least some of the intrinsic properties of things are somehow experiential, analogous to experience, or proto-experiential, but certainly not simply neutral. The Kantian point that such a theory speculates beyond the realm of possible experience is nevertheless well taken. If there are phenomenal properties in nature outside of our own consciousness, then we will never be able to access them directly. Panpsychism vastly expands the problem of other minds. By the same token, if panpsychism is true, then nature is much more similar to us. The conscious human mind is not an alien subject in a mechanistic material universe of Cartesian extended objects. If our

own conscious experience tells us—if only by analogy—something about the deepest levels of the universe, then the hiatus between mind and world may be less deep than modern philosophy has traditionally assumed. The nature of the thing in itself is not completely hidden from us.

We can conclude from these considerations that panpsychism is indeed a robust metaphysical position that is conceptually different from the neutral monisms inspired by the inscrutability of the ultimate nature of things.

2.3 Constitutive Panpsychism, Emergence, and Mental Causation

Constitutive panpsychism, sometimes also labeled 'compositional panpsychism,' is probably the most-discussed position in the current debate. It claims that macroexperience is constituted by, grounded in, or realized by microexperience. In its most attractive form it claims that there is an a priori entailment from microphenomenal truths to macrophenomenal truths. It seems to hit the sweet spot between physicalism and dualism. Due to its type-A a priori character, this position is a stronger modal thesis than the so-called type-B physicalism. It is endowed with all the explanatory power one could ask for. The only sort of identification which the type-B physicalist allows between physical and mental states or events is a posteriori. There is no a priori entailment of phenomenal concepts in physical concepts. In other words: Zombie worlds are perfectly conceivable but they are metaphysically impossible. Type-B physicalism thus rests on a form of modal dualism. There is a clear distinction between conceptual or logical possibility and real or metaphysical possibility. The realm of what is accessible to rationality by logical and conceptual analysis is disconnected from the realm of being, that is, from the realm of what is metaphysically possible. This gap undermines the entire project of employing conceptual analysis to do metaphysics. Metaphysical reasoning is then bound by what is accessible to us by sense experience. Science, not a priori analysis, discovers the nature of things.

In the case of the mind-body problem this leads to a kind of bruteness and opacity of the identity between the mental and the physical. It is in no way transparent to us how a relational structure described by physics necessitates conscious experience. We can discover these psycho-physical relations, but we have no conceptual insight into the nature of the necessity. It is this lack of analysis that motivates panpsychism. From a panpsychist point of view, a fully transparent analysis of the physical concepts will reveal why the basic physical entities will necessitate higher-level consciousness if they are appropriately arranged to promote a highly integrated flow of information. A fully transparent

analysis of the physical concepts will reveal the true nature of the physical, which includes mental or proto-mental intrinsic properties. For the type-B physicalist the most appropriate reply seems to be to claim that physical concepts are fully transparent but that mental ones are not. Talking about phenomenal consciousness does not reveal the true nature of consciousness. The true nature of, say, pain is not captured by the description of what it is like to be in pain. There are two reasons why this answer fails to satisfy the panpsychist. Firstly, for something to be pain, it has to feel like pain. Trying to define the identity conditions for pain without referring to phenomenal concepts seems like dodging the question about the nature of pain, not answering it. The other reason for doubting the type-B physicalist's strategy is given by the scientific principle according to which we understand things by breaking them down into smaller components. We wish to understand how the phenomenal mind, as we know it from human experience, is constituted by the material described by neurophysiology and physics. Type-B physicalism simply cannot provide a satisfying answer to the constitution question, unless it somehow gets rid of the phenomenal character of experience. The constitutive panpsychist, by contrast, provides an elegant answer to this question. To illustrate this it might be best to return to the so-called cellular automata mentioned above.

2.3.1 Cellular Automata

Gregg Rosenberg developed an argument against physicalism based on the idea of cellular automata (see Rosenberg 2004, 14–30). It requires fewer presuppositions than the zombie argument (like possible worlds, 2D semantics) but is nevertheless fully adequate to express the trouble with physicalism and pinpoint the solution suggested by the panpsychist. Cellular automata are artificial digital worlds consisting of basic particulars called 'cells' in an abstract space. These cells have relational properties connecting them to other cells. Computer modelers define various worlds by giving the cells different properties and then studying their dynamics through consecutive computational steps. This is usually done by defining rules that determine which properties a cell will have at a given time as a function of which properties the neighboring cells had at an immediately preceding time. In simple versions of cellular automata the basic particulars only have simple properties like 'on' and 'off.'

One can build on these humble beginnings and construct more complicated cellular automata that may ultimately mimic physical properties like spin, charge, or mass. The fascinating fact is that despite its rather simple physics, the cellular automaton is enormously versatile, in fact a universal Turing machine. Individual cells join together very quickly to build ever more complex structures and patterns that are sustained over many steps of computation. The machine seemingly produces endurants (i.e., stable relational patterns) that arise out of a sequence of event-like occurrents (i.e., discrete computational states of the system). These patterns become quite sophisticated, featuring a kind of nontrivial self-replication that is functionally similar to certain structures of living beings (i.e., DNA). For this reason cellular automata are sometimes called 'life worlds.' The basic facts of cellular automata—that is, the ways in which properties are distributed over the grid of cells—necessitate all higher-level structural facts about stable emerging patterns. There is no mysterious strong emergence involved, even though the emerging patterns exhibit new properties that cannot be attributed to individual cells. Everything happens within one strictly delineated ontological scheme. Physicalism can be construed as the thesis that our world is an extremely complex cellular automaton. But then the following problem arises:

- (1) The fundamental facts of cellular automata are defined entirely by the dynamic relations among the cells.
- (2) Facts of phenomenal consciousness are intrinsic qualitative facts, which cannot be entirely defined by the dynamic relations in which they enter.
- (3) Facts about dynamic relations do not entail intrinsic qualitative facts about phenomenal experience—neither a priori nor a posteriori.
- (4) Thus, the intrinsic qualitative facts about phenomenal experience are not entailed in the facts about cellular automata.

The question is then: If our world is a cellular automaton, how does the phenomenal mind emerge? The answer is straightforward for the constitutive panpsychist: The individual cells have intrinsic natures that are mental or at least analogous to mentality (that is, proto-mental). It is the composition of these intrinsic natures that explains the emergence of phenomenal minds. The emergence is thus neither brute nor inexplicable. The composition of the cells alone accounts for the weak emergence of higher-level structures and higherlevel mentality. There is logical synchronic supervenience between the lower and the higher levels. A perfect copy of all of the cells, including their intrinsic natures, will necessitate higher-level structure and higher-level phenomenal properties. The microdeterministic layered ontological framework of physicalism can be fully retained. The physical level determines all the facts, if 'physical' is taken in the broad sense such that quiddities, that is, intrinsic natures, are included. The beauty of constitutive panpsychism lies precisely in its ability to leave the overall framework of traditional physicalism intact. The macrofacts are synchronically microdetermined. Reductive explanations in the sciences are metaphysically vindicated. Constitutive panpsychism just adds nonobservable intrinsic natures to the scientific image. These natures do much

of the metaphysical heavy lifting in the philosophy of mind, without getting in the way elsewhere by interfering with the physical laws governing observable physical processes. For all pragmatic or instrumental scientific purposes it is perfectly acceptable to abstract away from those intrinsic natures. From the point of science the quiddities are mere metaphysical postulates. Thus even if constitutive panpsychism is true, science can work under the presumption of traditional methodological physicalism. Most significantly, constitutive panpsychism can hold on to the causal closure of the physical. David Chalmers sees this as a distinct advantage of constitutive panpsychism over its cousin, nonconstitutive panpsychism. The latter position requires strong emergence and (possibly) downward-causation and seems therefore prima facie incapable of providing a clear theoretical advantage over (emergent) dualism. But how, exactly, does constitutive panpsychism preserve the causal efficacy of the mental, especially in cases where beliefs or desires cause the movement of bodies?

2.3.2 The Causal Efficacy of the Phenomenal Mind

For the constitutive panpsychist, the causal efficacy of phenomenal properties does not rely on their being directly involved in causal relations. Rather, they are thought to be efficacious because they are essential properties of the entities that do enter into causal relations. Without some kind of intrinsic properties, the entire network of causal relations could not exist. This move is somewhat reminiscent of Davidson's theory of mental causation: Mental properties do not enter into the causal laws, but that-according to Davidson-does not render the mental epiphenomenal (see e.g., Davidson 1993). It is not the event qua mental or qua physical properties that is causally efficacious; it is the entire event, the event as such. Changing the mental properties will yield a different event and thus a different causal story. But doesn't this leave the phenomenal properties as being epiphenomenal? The entire relational network seems unaffected by the intrinsic properties. If one removed the intrinsic phenomenal properties and replaced them with some other intrinsic property, how would this change the physical causal network of causal relations between the events? The Davidsonian move might suggest something like this: The causal relation holds between individual events (not between types of events), and all properties of the event comprise its individual essence. Then causal relations could be affected by changing even seemingly irrelevant properties of a causing event. In the case of panpsychism: If the cause had different intrinsic properties, then the effect would be different. Thus the intrinsic properties are causally efficacious after all.

Sosa criticized this move with a now well-known example: A loud shot kills someone (Sosa 1984, 277). The loudness is epiphenomenal with regard to the killing. Had the gun been equipped with a silencer, the shot would have been equally fatal. In the same sense, mental properties are causally irrelevant within the Davidsonian framework. Davidson replied as follows: "Had the gun been equipped with a silencer, a quiet shot, if aimed as the fatal shot was, and otherwise relevantly similar, would no doubt have resulted in a death. But it would not have been the same shot as the fatal shot, nor could the death it caused have been the same death" (Davidson 1993, 17). A panpsychist might indeed claim that in a world with phenomenal or protophenomenal intrinsic natures, the relata of the causal relations are different from those in a world without such intrinsic natures. Something else is doing the causing and the effect is a different one; therefore the intrinsic properties are causally efficacious after all. But are the causal relations and the causal laws affected by a difference in intrinsic nature? It seems not. In an alternative zombie-world in which nonexperiential intrinsic properties carry the network of causal relations, this network is, on the assumption of the thought experiment, an isomorphic and indistinguishable copy of the causal network in our world. In this world something else is doing the causing, but the causal laws are the same. How are the assumed intrinsic properties of our world causally efficacious? They play a metaphysical role in carrying the causal relations, but they do not really determine those relations. Unless the metaphysical nature of the relata determines the causal relations themselves, the causal efficacy of intrinsic natures remains dubious. They are epiphenomenal with regard to the causal network and the causal laws.

2.3.3 Supervenient Causation?

The strong supervenience relation posited by the constitutive panpsychist could possibly do some work to secure the causal relevance of the mental. In the metaphysical framework of constitutive panpsychism, all causal relations at the macrolevel supervene logically on physical relations at the microlevel. Physical causation relates physical entities. If there are intrinsic properties of the physical, then it is instantiations of these properties that are related by physical causation. The intrinsic properties are mental in some sense (that is, they are microphenomenal). Because the phenomenal properties of the human agent supervene strongly on those microphenomenal properties, the former properties inherit the causal relevance of the latter (Chalmers 1996, 154). But what exactly is the causal relevance of these intrinsic properties? They are certainly not needed in scientific causal explanations. Causal laws

operate with physical properties only. Science abstracts away from the quiddities. Nonreductive physicalism is sometimes seen as arguing for dualism with regard to a duality of useful conceptual frameworks but for monism in metaphysics: Even though reality ultimately consists exclusively of physical entities, dualistic talk about the causal role of propositional attitudes and even of phenomenal properties is pragmatically indispensable. Constitutive panpsychism argues in a parallel, but inverted way: Even though reality ultimately has an intrinsic/extrinsic and thus mental/physical duality, it is pragmatically indispensible for science to limit itself to a monism which focuses on the relational physical structure only. But then there is no causal relevance for those intrinsic properties in our practice of explanation. If we abstract from the intrinsic properties, we have at the basic level a net of causal relations which are all physical in the narrow sense. The quiddities are relevant 'only' in a metaphysical sense.

But there are metaphysical worries as well. Does the concept of 'supervenient causation' really make sense? All higher-level causal relations in the world supervene logically on the basic physical relations. What could it mean that the former 'inherit' causal efficacy? The higher-level causal relations are completely microdetermined. All the causal work occurs at the base level, and the higher levels continue to enjoy a metaphysical free lunch. The higher levels qua higher levels have no causal efficacy whatsoever. The concept of 'supervenient causation' is misleading, since the supervenient level is asymmetrically dependent on the subvenient base. The causal work is completed at the subvenient level, leaving no causal work to be done by the supervenient level. 'Inheritance' is a misleading concept here, suggesting that the higher level inherits some good from the lower level, with which it could then do something. The supervenient level does nothing at all. It is completely and asymmetrically dependent. The case of logical supervenience might be special. Logical supervenience (across all possible worlds) is almost as strong as identity. A typeidentity between higher-level causal structures and the basic physical causal relations would suffice to carry the causal power all the way up. Identity is symmetrical, after all. If the macrolevel is identical with the microlevel, then the levels cannot compete for causal efficacy because no two things that are identical can compete for anything. Analogously, the strong modal force of the relation of strong supervenience might carry causal efficacy all the way up. But mental causation, as many understand it, should make sense of us as agents. If all the causal power is already located in the microconstituents, is there any room left for a robust sense of agency?

But even if we granted a monopoly on causal power to the microconstituents, the problem for the constitutive panpsychist does not disappear entirely. If the analysis of mental causation requires an account of the causal role of mental properties, then the Davidsonian move is ultimately futile. The basic structure of the problem is this (where 'physical' is used in the narrow sense):

- (1) Phenomenal facts do not logically supervene on physical facts.
- (2) Facts about causal relations do logically supervene on the physical facts.
- (3) Thus phenomenal facts cannot be constitutive for causal relations.

The only way out, it seems, would be to allow for mental-to-mental causal relations. Imagine that, in a cellular automaton, there exist causal laws of the following kind:

If in a given structural pattern, say, five adjacent cells have the intrinsic property of the type 'hot,' then in the next step the intrinsic property 'pain' will be activated in some cells. There are causal relations from intrinsic states of the cells to future intrinsic states. The relational properties of the cells remain unaffected by these changes.

The problem is that these changes are disconnected from the relational structure of the cellular grid. Since they have no effects on the relational structure, they form a separate and independent causal realm. In effect, this position will have to advocate some kind of psycho-physical parallelism that is somewhat reminiscent of Leibniz's position. But how could such parallelism preserve a robust account of agency in a physical world? What we wish to preserve and account for in any theory of mental causation is not primarily an abstract metaphysics of the causal efficacy of mental properties. It is rather our strong common-sense intuition that agency requires the causal efficacy of a higherlevel entity as such. Agency requires that my actions as a person are the result of me as a higher-level unity. If the microparticles that constitute me do all the causal work, then the sense of agency is lost. Constitution cannot account for the causal efficacy of the entities constituted at a higher level.

The question is whether the nonconstitutive panpsychist has a better answer to this question. A key problem for the constitutive panpsychist is the ontological status of higher-level mental unities. The panpsychist conceives of the human mind, for example, as constituted by a large number of smaller minds. The so-called combination problem surfaces as a 'composition problem.' Why does an appropriate arrangement of those smaller minds necessitate these higher-level unities? This seems to be a case of strong emergence, since it is conceivable that there is another possible world with the same microlevel mental entities in which no macrolevel mental entities supervene. Constitutive panpsychism must exclude this possibility. The higher-level mind must supervene logically on the appropriately configured microlevel minds. We have no clue
what the compositional principles are that allow this seemingly magical binding of experiences into larger units. We do not understand how subjects sum. This makes us inclined to believe that this special composition question can only be answered by resorting to some form of strong emergence. But the fact that we do not (yet) understand this does not mean that it might not in fact be how nature works. This is how it works everywhere else, the constitutive panpsychist claims. Isn't it then reasonable to assume that the mental as a natural phenomenon is no exception (that is, to assume that nature is homogenous)?

But if this much is granted, another problem looms. According to the constitutive panpsychist, the most basic entities—the furniture of the universe are experiencers endowed with a point of view or some form of subjectivity. Everything else is made up of these entities. But if this is so, then animals or human experiencers are also basic entities, because they are clearly experiencers and subjects in that sense. Yet the idea of the constitutive panpsychist was that the irreducibly basic entities are only found at the microlevel of nature and that everything else is a composite of them. But if higher-level subjects emerge over the course of evolutionary development, then they belong by definition to the irreducibly basic entities of the universe. Similarly, a proponent of substance ontology can claim that only the basic particles at the microlevel are the true substances making up the furniture of the universe: Carving nature at its joints means describing it at this level. Everything else is just a configuration of these building blocks. In this case a human being is not a substance but a mere configuration of substances. If, however, human beings are counted as substances, then they are ontologically quite independent and not mere configurations of smaller substances. This means that they must be counted as primitive nonreducible particulars in the furniture of the universe. This thought can also be applied *mutatis mutandis* to constitutive panpsychism. The unity of a higher-level subject of experience entitles it to be counted among the basic entities of the universe simply by being a subject of experience. It can then avoid being construed as a mere composition of smaller subjects of experience.

2.4 Nonconstitutive Panpsychism, Emergence, and Mental Causation

The nonconstitutive panpsychist bites this bullet. The special composition question is solved by assuming the strong emergence of higher-level mental unities, if the structural conditions at the lower level are met. But this move seems to siphon the theoretical elegance and beauty out of panpsychism. If something like the human mind (or some other animal's mind) does not logically supervene on microphenomenal properties, then we might as well be emergent dualists. Emergent dualism requires the strong emergence of the phenomenal from the nonphenomenal. Nonconstitutive panpsychism requires the strong emergence of macrophenomenal subjects of experience from microphenomenal subjects of experience. Wasn't the very motive of panpsychism to avoid such a strong emergence? Constitutive panpsychism requires only weak emergence. That is the crucial question that must ultimately be answered. But before we get there, we must first provide a sketch of the metaphysical picture of nonconstitutive panpsychism.

The key question, of course, is which notion of strong emergence the nonconstitutive panpsychist will employ. Even strong emergence comes in different flavors. The main difference in construing emergence seems to be between theories based on nomological supervenience (see Kim 1999) and theories which regard emergence as a nonsupervening causal relation (see O'Connor 2000). The latter is usually taken to provide a more solid grounding for higher-level individuals with novel causal powers (i.e. downward causation). Supervenience construed as an asymmetrical dependency of the higher level on the lower ones seems to preclude any real causal efficacy on the part of the higher levels. The second notion, emergence as nonsupervening causal relation, turns out to entail an even stronger emergence claim. It does this by breaking the bond of synchronic supervenience. If the nonconstitutive panpsychist takes this latter route, then the position will eventually collapse into emergent dualism—or so it seems. But a crucial and decisive difference will remain between the emergent dualist and the constitutive panpsychist. The emergent dualist can solve neither the ultimate-carrier problem nor the problem of the emergence of the phenomenal mind out of entirely nonphenomenal constituents. Nonconstitutive panpsychism, by contrast, has an answer to both of these questions and is thus clearly conceptually different from emergent dualism. Thus, even the version of panpsychism that is closest to classical dualism does not collapse into this well-established position but remains a genuinely distinct account in the philosophy of mind. In the remaining sections this account will briefly be sketched out.

2.4.1 Causal Supervenience

O'Connor and Wong argue that, if emergent entities are metaphysically primitive rather than mere constitutive resultants of lower-level features, then the correct relation between the lower and higher levels is causal, not supervenient (O'Connor and Wong 2005). If these causal connections are indeterministic then we can escape the grip of microdeterminism. For in this case, fixing the

microphysical state of the universe will not suffice to fix the distribution of emergent properties—even if we take emergence laws into account. If these laws are indeterministic, then a given microphysical state may have more than one emergent outcome. Adding temporal dynamics to this picture yields a causal process-oriented metaphysics instead of a formal and static view. The importance of this move cannot be overestimated. A historical example for it would again be Whitehead's "Process and Reality" (Whitehead PR). His classic statement "The many become one and are increased by one" (Whitehead PR, 32) captures the idea succinctly. If the underlying levels of nature reach a certain threshold of complexity of configuration, then an emergent individual is likely to appear. The causal properties of the emergent entity will go beyond the summation of the causal powers at the underlying microlevel. There will be a genuinely new entity with new causal powers. Can this account of emergence handle Kim's worries about causal exclusion (see Kim 1999)? It seems so. On Kim's view, the diachronic causal activities of an emergent entity are metaphysically superfluous. They add no causal efficacy over and above that of the causal mechanisms at the base level. But this contention is based on the asymmetric dependence built into the concept of supervenience. The O'Connor/ Wong model is not susceptible to this kind of counterargument. The entities at the basic or subvenient level do not determine the emergent effects independently of the causal activities of those emergent entities (see O'Connor and Wong 2005, 670).

The idea can again be expressed by the analogy of the cellular automaton or 'life world.' Imagine a very complex and long-lasting three-dimensional life world, which mirrors the complexity of our universe. For the longest time, the weak emergence of higher-level structures occurs exactly in accord with the rules governing the individual cells. But occasionally a macroobject appears whose behavior diverges slightly from what the rules predict. The events at the microlevel are affected by this macroobject, changing whenever and wherever such a higher-level structure is present. We would then be able to assign new rules to the situations in which these higher-level structures emerge. It follows from this that the low-level rules do not necessitate the entire future dynamic of the system, not even at the lowest of the hierarchical levels of our three-dimensional cellular automaton. No knowledge of the behavior of those macroobjects or of their effects on the entire system can be derived from knowledge of the smallest microobjects and the rules they follow. If the basic rules are indeterministic such top-down influence might even happen without 'breaking' the most fundamental rules, only the probability distribution will be slightly affected. Such a system is conceptually coherent. If our world were like this, then it would contain strongly emergent entities with downward causal powers. An alternative nonsupervenience view of emergence that would

allow for genuine causal efficacy of emergent entities is the fusion account of emergence. Fused entities cease to exist as separate entities. The emergents created by fusion are endowed with new causal powers because the fused entities lose some of their original causal powers (see Humphreys 1997).

2.4.2 Is This Compatible with Science?

Even if these accounts of nonsupervenient emergence are certainly a conceptual possibility, many will claim that we lack empirical evidence for emergent causation. But there are at least candidates for such causal emergence. In mathematical simulations of neural systems it can be shown that the macrolevel can causally supersede the microlevel. This causal independence of the macroobjects, which are more than the sum of the underlying microobjects, is indeed a key feature of Tononi's "integrated information" theory of consciousness (see Hoel, Albantakis and Tononi 2013). Another example may arise in quantum mechanics with the possibility of emerging holistic properties (see Schaffer 2010). Quantum entanglement may be a case in question. Prosser's elegant recent account of this idea can serve as an illustration (see Prosser 2011). Physics does not know of any viable procedure for reducing the entangled state to a summation of classical states and hence reducing quantum mechanics to classical physics. But if that is granted, then the properties of entangled atoms might well be causally efficacious for the future dynamics of the world. Prosser argues: "entanglement shifts a probability distribution concerning the behavior of atoms-the overall configuration of the . . . atoms is likely to be different when there is entanglement—and its effects therefore constitute downward causation. Hence there is no conflict between downward causation by the emergent property of entanglement and the base-level laws" (Prosser 2011, 37). Humphreys's idea of fusion emergence explicitly regards quantum entanglement as involving a fusion of entities into unities of a new kind (see Humphreys 1997). Whether strong emergence in this sense exists is an empirical question. Yet even if the above examples are empirically inadequate, other ways of introducing strong emergence might be found. The mysterious collapse of the wave function itself might be another promising starting point.

2.4.3 Why Nonconstitutive Panpsychism Differs from Emergent Dualism

But why would the panpsychist want to be a strong emergentist? Isn't panpsychism's greatest advantage that it can avoid strong emergence? If higher-level individuals with new causal powers can strongly emerge, then why can't phenomenal minds strongly emerge from an entirely nonphenomenal mindless world? If strongly emergent panpsychism is possible, then so is strongly emergent dualism. And dualism seems intuitively more appealing, since it does not commit us to externally unobservable quiddities or microphenomenal intrinsic properties at the lowest level of nature.

Panpsychism rests ultimately on two theoretical claims: the argument from intrinsic natures and the genetic argument. The argument from intrinsic natures claims that the relational structure described by physics is incomplete. This structure stands in need of categorical intrinsic natures which can carry the relational network of physical properties. As noted above, Leibniz made this point against Descartes, and in the twentieth century Russell and Whitehead advanced the same argument. The dualist has no adequate reply to this objection. Dualism lacks a metaphysically plausible theory of matter. Idealism, which deflates matter to a 'well-founded phenomenon' (see Leibniz), escapes this problem. But the Cartesian substance dualist faces not just the well-known problems of causal interaction between spatial and nonspatial entities, but also that of developing a plausible metaphysical theory of matter. Mere extension or mere relation will not do the trick. It leaves open the questions of what is being extended and what is being related. Panpsychism overcomes this dearth of analysis. This is the first reason why nonconstitutive panpsychism does not collapse into emergent dualism.

The second reason is related to the genetic argument for panpsychism. Even O'Connor and Wong, who are no panpsychists, admit that their concept of emergence requires a "tendency had by each of the basic entities" (O'Connor and Wong 2005, 665) which explains the strong emergence of phenomenal minds. But what exactly does "tendency" mean here? Can something entirely nonphenomenal and nonmental have the tendency to bring about phenomenal minds? Earlier we introduced the distinction between inter-attribute and intraattribute emergence. Inter-attribute emergence might be labeled 'superstrong emergence.' In addition to weak and strong emergence we thus introduce superstrong emergence. Weak emergence (based on strong-supervenience relations) and strong emergence (based on weak-supervenience or causal relations) occurs within a unified categorical framework. For example: Higherlevel spatio-temporal concrete entities can only emerge either weakly or strongly from lower-level spatio-temporal concrete entities. Superstrong emergence breaches or transcends categorical frameworks. A clear case of superstrong emergence would arise if something emerged from absolutely nothing. Another clear case of superstrong emergence would arise if a concrete spatiotemporal entity emerged in a world in which only abstract entities exist. The panpsychist claims that another case of superstrong emergence is given by the

emergence of the phenomenal mind from a world which is merely spatial extension or a framework of causal-functional relations.

The nonconstitutive panpsychist needs strong but not superstrong emergence. There are new unified entities endowed with phenomenal minds that are more than just the constitutive summation of smaller such entities. But even the smallest entities have some form of phenomenal properties. Phenomenality was there in the beginning. Likewise, the strongly emerging higher-level entities feature new causal powers that are more than just the constitutive summation of the causal powers of smaller such entities. But even the smallest entities have some form of causal properties; causality was there in the beginning. What is new is simply the strong emergence of new natural individuals (concrete entities) that cannot be reduced to their constituents. What strong emergence makes possible is the 'special composition' of new entities from the same basic makeup as their constituents. Strong emergence cannot, however, create something absolutely new in its metaphysical nature. This recalls again to mind Whitehead's classic dictum: "The many become one and are increased by one" (Whitehead PR, 32). Emergent dualism postulates the emergence of a whole new metaphysical category of entities. Mental entities emerge from nonmental entities. This is a case of superstrong emergence as defined above. Emergent dualists thus not only have no answer to the problem of causal pairing and that of the intrinsic natures, they also require superstrong emergence for the mental to emerge from the physical. Thus, panpsychism, even in the strong-emergentist form of nonconstitutive panpsychism, clearly differs from dualism.

Can the nonconstitutive panpsychist get a grip on the problem of mental causation? The O'Connor/Wong model of emergence contains genuine downward causation. Higher-level entities endowed with minds could in principle have causal powers that are not microdetermined by the causal relations at lower levels. The nonconstitutive panpsychist therefore need not be burdened with the problem that all macrocausal relations strongly supervene on microcausal relations. There is an opening for macrolevel agents to make a causal difference in the world. This is a genuine difference from constitutive panpsychism.

Some would argue, however, that genuine mental causation requires that the mental content of the psychological states of these entities or agents be causally efficacious. The nonconstitutive panpsychist can block this argument by the very 'Davidsonian move' made above by the constitutive panpsychist. According to the Davidsonian move, it is the entire entity that is doing the causing, and if its intrinsic (mental) properties were to change, something else would do the causing and something else would be caused. But again, this option renders too many properties causally efficacious, as Sosa rightly points out (Sosa 1993). The problem of mental causation seems all but intractable. No position in the philosophy of mind has a fully convincing answer to it. But nonconstitutive panpsychism is certainly not in a theoretically precarious position with regard to mental causation. It has the theoretical advantage over emergent dualism that it does not require a causal pairing of spatial and nonspatial entities. Nonconstitutive panpsychism is again clearly distinct from dualism.

2.5 Taking Stock

It has been shown that panpsychism is a conceptually stable and independent position in the philosophy of mind. It differs from physicalism, dualism, idealism, and even neutral monism. It comes in two flavors. One—constitutive panpsychism—incorporates many ideas from type-A physicalism without collapsing into it. The other—nonconstitutive panpsychism—incorporates ideas from emergent dualism without collapsing into it. I admit to having (at least sometimes) some sympathies with Kantian worries about the metaphysical inscrutability of the relation between the mental and the physical. So I am not claiming that a version of panpsychism is true. But I am claiming that it might be. Both versions of panpsychism needs to be taken seriously. Thus: "If it is not true, it is well conceived—Se non è vero, è molto ben trovato."

Bibliography

- Brüntrup, Godehard. "Natural Individuals and Intrinsic Properties." Unity and Time in Metaphysics. Ed. Ludger Honnefelder, Edmund Runggaldier, and Benedikt Schick. Berlin/New York: de Gruyter, 2009. 237–52.
- Brüntrup, Godehard. "Panpsychism and Structural Realism." *New Perspectives on Panpsychism*. Ed. Michael Blamauer. Frankfurt: Ontos, 2011. 15–35.
- Chalmers, David. The Conscious Mind: In Search of a Fundamental Theory. Oxford: Oxford University Press, 1996.
- Davidson, Donald. "Thinking Causes." *Mental Causation*. Ed. John Heil and Alfred Mele. Oxford: Oxford University Press, 1993. 3–18.
- Haugeland, John. "Pattern and Being." *Dennett and his Critics*. Ed. Bo Dahlbom. Oxford: Blackwell, 1993. 53-69.
- Hoel, Erik, Larissa Albantakis, and Giulio Tononi. "Quantifying Causal Emergence Shows that Macro Can Beat Micro." *Proceedings of the National Academy of Sciences* 110.49 (2013): 19790–795.
- Humphreys, Paul. "How Properties Emerge," Philosophy of Science 64.1 (1997): 1-17.
- James, William. The Principles of Psychology. Cambridge: Harvard University Press, 1983.
- Kant, Immanuel. Critique of Pure Reason. Transl. Norman Kemp-Smith. London: Macmillan, 1929.
- Kim, Jaegwon. "Making Sense of Emergence." Philosophical Studies 95.1-2 (1999): 3-36.

Leibniz, Gottfried. *Die philosophischen Schriften*. Vol. I–VII. Hildesheim: Georg Olms, 1965. Newman, Max. "Mr. Russell's Causal Theory of Perception." *Mind* 37.146 (1928): 137–48.

- O'Connor, Timothy. Persons and Causes: The Metaphysics of Free Will. Oxford: Oxford University Press, 2000.
- O'Connor, Timothy and Hong Yu Wong. "The Metaphysics of Emergence." *Nous* 39.4 (2005): 658–78.
- Prosser, Simon. "Emergent Causation." Philosophical Studies 159.1 (2011): 21-39.
- Putnam, Hilary. "Models and Reality." Journal of Symbolic Logic 45.3 (1980): 464–82.
- Rosenberg, Gregg. A Place for Consciousness. Probing the Deep Structure of the Natural World. Oxford: Oxford University Press, 2004.
- Russell, Bertrand. Our Knowledge of the External World as a Field for Scientific Method in Philosophy. Chicago: Open Court, 1914.
- Russell, Bertrand. The Analysis of Matter. London: Routledge, 1927.
- Schaffer, Jonathan. " Monism. The Priority of the Whole." *Philosophical Review* 119.1 (2010): 31–76.
- Sosa, Ernest. "Mind-Body Interaction and Supervenient Causation." *Midwest Studies in Philosophy* 9.1 (1984): 271–81.
- Sosa, Ernest. "Davidson's Thinking Causes." *Mental Causation*. Ed. John Heil and Alfred Mele. Oxford: Oxford University Press, 1993.
- Whitehead, Alfred. *Process and Reality: An Essay in Cosmology*. Corr. ed. Ed. David Griffin and Donald Sherburne. New York: Free Press, 1978.

PART II

THE VARIETIES OF PANPSYCHISTIC ONTOLOGIES

3

Mind and Being

The Primacy of Panpsychism

GALEN STRAWSON

3.1 Introduction

I'll start with a metaphysical creed—four propositions.¹ I'm confident that the first three are true, and I suspect that the fourth is true, but I don't think one has to accept any of them to agree with my principal thesis—the thesis of *the primacy of panpsychism*, the highly unoriginal thesis that there are compelling reasons for favoring panpsychism above all other positive substantive proposals about the fundamental nature of concrete reality.²

I'll state the four propositions first in German because I like the way they sound in German. (1) *Stoff ist Kraft*, (2) *Wesen ist Werden*, (3) *Sein ist Sosein*, (4) *Ansichsein ist Fürsichsein*. These are identity claims—fully reversible. I'm not going to argue for them, but I'll provide a few glosses.

3.2 Stoff ist Kraft

(1) Stoff ist Kraft.

Matter is force, or as I will say energy:

(1) *matter is energy*.

Strictly speaking matter is only one form of concrete being, but I'll use the word loosely to mean all concrete stuff: all concrete being is energy—energy-activity, energy-stuff.

I'm using the word 'energy' as Heisenberg does when he writes that "energy is a substance," "all particles are made of the same substance: energy" (Heisenberg 1958, 63; 71); putting aside the common use according to which 'energy' denotes the power of 'doing work' contained in or possessed by a body or system of bodies.

I take this general position to be orthodoxy today. It's also an old view, if Aristotelian *energeia* can be understood as energy: "in Aristotle the concept *energeia* coincides with that of reality. And Leibniz, too, declared: *'quod non agit, non existit'*" (Schlick 1918–25, 181): what doesn't act doesn't exist. The most fundamental characterization of substance is *that which acts*: "activity . . . is of the essence of substance" (Leibniz 1714, 65).³

Some may find the equation of force, energy, power, and activity too quick or easy. I think it's eminently defensible—the evolution of the old notions *dunamis* and *energeia* into the notions of dynamism and energy is highly significant. I also take it that the existence of *causation*—of the "because something is, something else must be" phenomenon as it concretely exists in nature⁴—is nothing over and above the existence of energy. One might say that the causal laws for our universe describe the particular form of energy as it exists in our universe—the behavioral form of energy as it exists in our universe.⁵

If spacetime is itself a concrete existent, something substantival, as I'm inclined to suppose along with many others, rather than a mere container for concrete being, and if it is in fact the only concrete existent, as a good number of physicists and cosmologists suppose, if in other words spacetime is the universe, if

(A) Sein ist Raumzeit,

if (A) *being is spacetime*—then, given (1), spacetime is energy.⁶ If the existence of spacetime is the existence of certain fields (electromagnetic, weak and strong nuclear, gravitational, Higgs), or ultimately only one field, then the existence of the field or fields is just a matter of the existence of energy.⁷

We may be very wrong about the nature of spacetime, insofar as our conception of spacetime goes in any way beyond our best equations—even if our best equations are essentially correct. So be it. I'll take the word 'spacetime' to be a name for the actual dimensionality of reality, the actual existence-dimension or *Existenzraum* of concrete reality *whatever its ultimate nature*—a term that leaves room for the possibility (the likelihood) that we are in certain ways bewilderingly wrong about it.

One point is worth noting straight away (I'll return to it later). We certainly shouldn't suppose that having spatial existence entails having some sort of irreducible nonexperiential stuff being, any more than we should suppose something we already know to be false (given that there is space): that having spatial existence is incompatible with experiential being.

I'm inclined to think that (1) is at bottom an a priori truth. Aristotle didn't wait for it to be presented as a scientific discovery. The once popular idea of inert or powerless concrete being is I believe incoherent,⁸ and the natural thought that powers require 'categorical grounds' doesn't require one to think that there is or can be any 'real distinction,' in Descartes's sense, between a thing's possession of the powers it possesses and the existence of those powers' categorical grounds, or that the existence of the categorical grounds can't be wholly a matter of the existence of energy. Granted that you can't have powers without 'categorical grounds,' so too you can't have categorical grounds without powers. Imagine an exhaustive specification of a thing x's powers P and categorical properties C (the specification of powers will be-benignlyinfinite if it's given in terms of x's possible effects on other things). It's plausible that only something identical to x in respect of C can possibly have precisely P and conversely that anything identical to x in respect of C must have precisely P. In this case neither C nor P can possibly exist apart from the other, so there's no real distinction between them in Descartes's sense, and where there's no real distinction between two things, it's plausible that they're really identical.

So much for the first proposition.

3.3 Wesen ist Werden

(2) Wesen ist Werden

i.e.

(2) *being is becoming.*

This is the essential (*Wesen*) nature of concrete being, of nature (*Wesen*). Everything is process, in other familiar terms. Being is process. Being is doing, activity. A through-and-through processual view of reality is mandatory. All concrete being is essentially time-being—whatever exactly time is. Being is be*ing*. All being is in Kant's phrase 'always already' behaving, becoming, and of course conversely (see Kant 1781–87, A346/B404).

(1) and (2) are close to

(5) Wirklich ist, was wirkt

—the actual is what has an effect.⁹ In the case of matter, Schopenhauer observes, "its being [*Daseyn*] is its acting [*Wirken*]: and it is inconceivable that matter has any other being."¹⁰ To say this is not to 'desubstantialize' matter in any way, or to 'operationalize' our conception of matter, and it is most emphatically not to suggest that matter is really only what we can possibly observe (as per the fatal modern tendency to epistemologize metaphysics—or metaphysicalize epistemology). It's simply to express in a certain way the point that the nature of concrete being is energy. The point is old, but we periodically lose hold of it. David Lewis has misled many with his extraordinary view (perhaps a legacy of positivistic empiricism) that the intrinsic or categorical nature of matter is or could be independent of its behavior.¹¹ (In what follows 'intrinsic' can usually be replaced by 'categorical'. A thing's categorical nature includes its power nature, since its powers are wholly a function of its categorical being.

So much very briefly for (2). There is of course a great deal more to be said about this.

3.4 Sein ist Sosein

(3) Sein ist Sosein.

This is harder to render in a single English sentence. I propose

(3) being is quality.

There's no metaphysically fundamental distinction between substance and attribute (as Descartes, Spinoza, Kant, Nietzsche, and many others agree). There's no metaphysically fundamental distinction between the concrete being of substance, 'thatness,' and the concrete being of (intrinsic) propertiedness, 'howness,' 'thusness,' qualitativity. There is no *real distinction*, in Descartes's terms, between a concrete entity's *Sein* and its *Sosein*, when that entity is considered at any particular time. There's no difference between bare being (the barest that being can get) and how-being: between being and being-some-way. Lewis is wrong to suppose that a thing's behavior (in any given context) could fail to be wholly a function of its intrinsic nature, for a thing's behavior in any given context is simply (and wholly) part of its being, its intrinsic nature. He is, however, right, on the present view, that concrete reality is "an arrangement of qualities. And that is all" (Lewis 1986, x).

It's obvious that there can't be *Sein* without *Sosein* or *Sosein* without *Sein*. To be at all is necessarily to be somehow, and to be somehow is necessarily to exist. The present stronger claim—that there's no real distinction, in the case of any

particular thing or object o, considered at any particular time t, between the totality of what constitutes the existence of o at t and the totality of what constitutes the existence of the (intrinsic instantiated) propertiedness of o at that time—may seem less obvious, but it's no less secure. Neither o at t nor o's (intrinsic instantiated) propertiedness at t can coherently be supposed to exist apart from the other in any respect at all, let the counterfactuals fall as they may. They are metaphysically identical—the same thing. We can express this as a subthesis of (3):

(3*) An object considered at any given time t = its (intrinsic/categorical instantiated concrete) propertiedness at t.

The way in which object words and property words operate in everyday thought means that this outright identity statement can sound plainly incorrect. In particular, the ease and naturalness with which we use counterfactual idioms when talking about objects and their properties can mislead us. We may for example be tempted to think that it is a sufficient objection to (3^*) to say something like: 'This very object considered now at *t* could have had different properties now at *t* from the properties it does in fact have.' In fact this is no objection to (3^*) , but it can take a certain amount of effort to rethink one's conception of the phenomena that lead us to talk of objects, on the one hand, and the phenomena that lead us to talk of instantiated properties, on the other hand, up to the point at which (3^*) no longer seems incorrect, but rather evidently true.¹²

Kant gets this exactly right, I think, when he says that "in their relation to substance, accidents [or properties] are not really subordinated to it, but are the mode of existing of the substance itself" (Kant 1781–87, A414/B441). There's no sort of ontic subordinacy of the object's properties to the object itself, no sort of existential inequality or priority or superiority or inferiority of any sort, no ontic dependence of either on the other, no independence of either from the other. In the case of any concrete entity, again, its *Sosein* (its being the way it is) is identical to its *Sein* (its being).

I take this claim to be a priori, however much language beguiles us to think otherwise. There is really no other possible relation of thatness to howness. But it is also perhaps the hardest of the four claims to grasp. Or rather it's the hardest to hold onto in such a way as to be able to deploy it properly in one's philosophical thinking. One can lose a theoretically live grasp of it when one isn't concentrating even if one endorses it whenever one focuses on it without trying to do anything else. I think this is principally because of our deep natural tendency to think of objects and their properties in counterfactual ways. These ways of thinking are perfectly in order, and crucial for many ordinary purposes, but they pull against (3) and (3*) in a way that can easily throw us off track in metaphysics.

A good way to avoid philosophical trouble caused by the words 'object' and 'property' is to use the maximally neutral word 'being,' which covers both of them. I'll sometimes do so. It's also important to see that (2) and (3) make it as legitimate to talk of experience or consciousness or experientiality as a kind of stuff as it is to talk of anything else—sugar, lead, matter, energy—as a kind of stuff. This is another considerable stumbling block, given current habits of thought; there's nothing much I can do about this. It takes time to habituate to the point, in spite of all the work that Russell, Whitehead and others put in, and in spite of the fading away of the idea that the being of phenomena like spin and charge requires the being of some further underlying distinct stuff to be the bearer of these properties. Perhaps the best thing to say is that there is certainly no more reason to think of matter as kind of stuff than there is to think of experientiality as a kind of stuff. So if one is comfortable with thinking of matter as a kind of stuff one should be—or needs to become—equally comfortable with thinking of experientiality in this way.

Certainly experience requires a subject of experience. But this doesn't mean we have to fall back to a metaphysics of object and property, with the subject of experience as object or substance possessing an ontologically distinct property: experience. What we have is the process phenomenon *subject-having-experience*; or—in other terms—the process-stuff *experience*, which is (necessarily) *subject-or-subjectivity-involving* process-stuff.

So much for the third proposition.

3.5 Ansichsein ist Fürsichsein

(4) Ansichsein ist Fürsichsein.

For something to be, to be at all, to be what it is considered wholly *in itself* or *an sich* in Kant's sense, is essentially for it to be *for itself*, in the familiar sense of this phrase according to which for a being to be 'for itself' is for there to be something it is like to be it, experientially. It is for it to be a conscious or experiencing being.¹³ So

(4) being is mind.

Being is essentially experience-involving—where 'experience' denotes any sort of conscious experience whatever, including the most primitive forms of 'mere' sensation. (4) is a form of panpsychism or panexperientialism (I use these terms interchangeably, taking the psychical to be essentially a matter of conscious—experiential—goings-on). It's equivalent to

(6) Sein ist Bewusstsein.

It's a form of *pure* panpsychism, which I here take to be the view that experientiality is all there is to the intrinsic nature of concrete reality (note that on this view, the existence of subjects of experience can't be supposed to be anything ontologically over and above the existence of experiencing). The milder version says that this is how things are in this universe. The stronger version says that this is all that being can be—that panpsychism is necessarily true.

I think this may be so—that concrete being has in fact no other possible form than energy, and that energy has in fact no other possible form than experientiality. But it certainly isn't a priori viewed from here. We can't hope to prove that the notion of nonexperiential concrete being is incoherent, even if it's a priori in God's physics, or rather his entirely general theory of being, his 'concretics,' his necessarily utterly comprehensive account of what can concretely exist. Still, the great William James holds that "our only intelligible notion of an object *in itself* is that it should be an object *for* itself," and that "a thing in itself . . . must be an experience *for* itself" (Perry 1935, 446). (See §21 below for a suggestion about why this might be so.)

Note that I'm understanding *experience* to entail *mind* (and conversely). One important terminological alternative is to define mind more narrowly, taking it that *mind* entails some sort of *intelligence*, and that experience as just defined—the stuff of reality, according to the present view—may exist without mind.¹⁴ This may be what Russell and William James have in mind when they propose that the basic stuff of reality is 'sensation' or 'pure experience' respectively, and nevertheless declare themselves to be 'neutral monists' who hold that the stuff of reality is neither mental nor physical. On my view their positions appear to classify as forms of panpsychism. But the disagreement is, so far, merely terminological.

3.6 The Basic Creed

So here's the basic creed: being is energy, process, quality, mind (experience). These four things are, in this universe, all the same thing—which may be spacetime (in which case there is no real distinction between concrete being and its *Existenzraum* or dimensionality). Once the restriction to concrete being is in place the four initial terms—*Stoff, Wesen, Sein, Ansichsein* come to the same thing: *Wirklichkeit*—the actual. That's the proposal. It's the backbone of the metaphysics I favor: *identity metaphysics*. The principal characteristic of identity metaphysics is that it finds identity where other metaphysical positions, dancing to the panpipes of language, find distinctness and difference. In the background stand powerful thinkers—Spinoza, James, Nietzsche among others, perhaps also Whitehead—although none, perhaps, would accept the whole of the basic creed unreservedly.

Perhaps I should say that I'm not arguing, so I'm not begging questions. I'm offering a picture of how things may be.

3.7 Natura Non Facit Saltum: No Radical Emergence

Let me now add a version of an old metaphysical thesis to the ontological theses (1)-(6):

(7) *natura non facit saltum*

i.e. (roughly) there are no absolute or radical qualitative discontinuities in nature.¹⁵ I take (7)—No Jumps—to be a solid part of any sound naturalism, and from (7), as I understand it, one can derive the No Radical Emergence thesis as I understand it, that is,

(8) there is no radical emergence

(some may think that (8) is effectively the same as (7)). And from (8), I submit, we can derive

(9) the experiential (experiential being) can't emerge from the wholly and utterly nonexperiential (wholly and utterly nonexperiential being)

—because any such emergence would have to be radical in the impossible way. I'm not going to argue for (8) and (9).¹⁶ The general idea is simple. Emergence—*e-mergence*, no less—can't be brute. In all genuine (nonradical) cases of emergence of one thing from another there's a fundamental sense in which the emergent phenomenon, say Y, is wholly dependent on—somehow wholly flows from—that which it emerges from, say X. Otherwise it simply won't be true after all to say that Y is emergent from X, for some part or aspect of Y will have come from somewhere else. (I understand emergence in what I take to be a standard way as paradigmatically a matter of constitution, not causation: in the present case, it would be a matter of individually nonexperiential phenomena coming to constitute experiential phenomena simply by coming together or being arranged in a certain way—as nonliquid H₂O molecules together come to constitute something liquid.) Many will agree. Others won't. Two things seem worth saying straight away. The first is that it's metaphysically far more extravagant and antinaturalistic to reject (7) the No Jumps thesis, and postulate radical emergence of the experiential from the nonexperiential, than it is to postulate nonradical emergence of the human or biological experiential from the nonhuman or nonbiological experiential—whatever difficulties the second idea may also seem to raise (e.g. the 'combination problem'—see §19 below).

Secondly, and more importantly, one doesn't need to meet those who don't agree with No Radical Emergence with an argument to support it. All one has to do is ask them politely why they think anything nonexperiential exists; especially when this belief forces them to endorse radical emergence, given that they're realists about experience.

On this more later. First, some more declarations.

3.8 Real Naturalism

Like many, I'm a *monist*, a *stuff monist*, an only one-kind-of-stuff monist:

(10) stuff/kind monism is true.

I'm putting aside only-one-thing monism, thing monism, according to which

(B) there is only one thing (object, entity, substance) in concrete reality,

for purposes of discussion, although, like many, I'm attracted to one version of it, that is, (A): the view that spacetime is a single thing—the universe.¹⁷

I'm not only a stuff monist. Like many again, I'm a *materialist* or *physicalist* monist (I use the words 'materialist' and 'physicalist' interchangeably)— someone who holds that everything that concretely exists in our universe is wholly physical:

(11) materialism/physicalism is true.

I also take it that everything that concretely exists is wholly *natural*—in no way supernatural or nonnatural. So I'm an outright ontological naturalist.¹⁸

I am however a real naturalist, a real materialist—unlike some who call themselves 'naturalists.' I don't disagree with them because they believe in the existence of something I judge to be supernatural. On the contrary: I disagree with them because, overtly or (more often) covertly, they doubt or deny the existence of a wholly natural concrete phenomenon we know to exist: the phenomenon of consciousness—conscious experience—experiential 'whatit's-likeness'—the phenomenological character of experience—the subjective qualitative character of experience. I understand all these five common phrases to denote the same thing, which I'll call 'experience,' instead of 'consciousness,' because the word 'consciousness' has been used in too many different ways.

So I'm an outright realist about experience, a real realist about experience:

(12) there is experiential concrete reality.¹⁹

Any real naturalist must be a real realist about experience, because experience is the most certainly known concretely existing general natural phenomenon, and is indeed the first thing any scientist encounters when they try to do science.

I say that I'm a *real* realist about experience because some who claim to be realists about experience aren't really any such thing. What do I mean by real realism about experience? The quickest way to say what it is is to say that it's to hold exactly the same general view about what experience is (color experience, say, or pain experience, or taste experience), considered specifically as experience, that one held *before one did any philosophy*, when one was thirteen or ten or six. One then had an entirely correct view. If people ask what that view is I'll ask them to think back to their childhood. If they say they still don't know I won't believe them.

So I'm a real naturalist and a real materialist—a materialist in the sense in which every materialist was a materialist until some time well into the twentieth century. That is, I'm someone who thinks that everything that exists is wholly physical and who is also fully realist about experience or consciousness. At the same time I know that 'physical' is a 'natural-kind' term, like 'gold,' or 'tiger,' and that we may be very ignorant (or plain wrong) about the fundamental nature of the physical in various ways—*if and insofar as* the fundamental nature of the physical is anything more than experience. So really the core meaning of 'physical' for me is just: 'concretely real.'

But in that case why do I say I'm a materialist? Because I believe that

(13) the claims of physics apply to everything that concretely exists

and also that

(14) many of the claims of physics are true of everything that concretely exists

(e.g. f = ma, the inverse square laws, etc.). And I also know something that was a philosophical commonplace in the early twentieth century, and indeed

earlier, and is fortunately becoming one again. I know that physics is "just a set of rules and equations," in Hawking's words (Hawking 1988, 174).²⁰ I know that

(15) physics can't characterize the intrinsic nonstructural nature of concrete reality in any respect at all

and a fortiori that

(16) physics has no terms with which to characterize the intrinsic *experiential-qualitative* nature of concrete reality,

whether only part of concrete reality has an experiential-qualitative nature, as we usually suppose, or whether all of it does, as panpsychists suppose. I know that physics is simply silent on the question of the intrinsic nonstructural nature of reality.

We ordinarily suppose that we have some positive nonstructural conception of the intrinsic nature of space or spacetime. So be it—so long as we're clear that this conception of space or spacetime goes beyond anything that the equations of physics tell us. One of the greatest difficulties that arise in the metaphysics of mind is precisely that we standardly and perhaps irrepressibly suppose that *physics* supports the accuracy of our basic imaginative picture (I mean 'imaginative' literally) of what spatiality is—and of what matter is. It doesn't.

So physics is silent about the intrinsic nonstructural nature of reality. The question is then this (it's an ancient question, but I'll give it again in Hawking's words): "What is it that breathes fire into the equations and makes a universe for them to describe?" (Hawking 1988, 174). What is it that the equations are true of? What is the fundamental, intrinsic, *nonstructural* nature or *stuff being* of the concrete reality that the true statements of physics are true of?

Call this fundamental intrinsic nonstructural stuff nature x. The relevant options are (i) x is wholly experiential (the pure panpsychist option), and (ii) x is partly experiential and partly nonexperiential (as most people suppose). I'm ignoring the third option, (iii) x is wholly nonexperiential, because it denies the existence of experience.²¹ The central claim of this paper is that (i) is the best option—that panpsychism is the most plausible theory of x, given a genuinely naturalist—materialist monist—outlook.

I'll now flag an assumption that is built into the question, and then make one more general metaphysical assumption specifically for the purposes of discussion.

3.9 Two More Assumptions

The assumption built into the question is that for any concrete entity x one can always distinguish between x's structural features and something about x that isn't just a matter of structure, something in virtue of which x has or exemplifies the structure it does, something that is therefore not itself just a matter of structure.

The assumption can seem very secure but it has been questioned. Ontic structural realists²² claim precisely that structure is all that concretely exists:

(C) concrete being is (wholly a matter of) structure.

And while Max Newman's claim that "it seems necessary to give up the 'structure–quality' division of knowledge in its strict form" (Newman 1928, 147) is an epistemological claim, it may be thought to point forward to an ontological proposal.

If one takes the structural properties of a concrete thing x to be properties that can be fully characterized in abstract, logico-mathematical terms, as I do, then I think one can safely conclude that (C) must be false. One can conclude that there must be more to x than merely its structural properties, on the seemingly secure ground that there must be more to concrete being than abstract being.²³ If, however, one understands structure in a richer way as something concrete—as 'causal structure' or 'spacetime structure' or (in a Schopenhauerian-Russellian fashion) 'spacetimecause structure'—then one may be able to link (C) to (A), and also—via the power-energy-causality equation—to the basic creed (1)–(5).²⁴

The further assumption that I'm going to make for the purposes of discussion is that

(D) there are a great many ultimate constituents of physical reality.

(D) is sometimes called 'smallism.'²⁵ It's very widely accepted, but—as is now clear—I'm not sure it's true. It would obviously be false if any version of (B) (thing monism) were true, and although it seems extraordinarily difficult to understand how any version of (B) could be true, given the seemingly evident and irreducible plurality of concrete things, it may yet be, as already remarked, that (A) there is a fundamental sense in which spacetime is indeed the only thing there is, and that all the particle phenomena recognized in the current standard model are just "various modes of vibration of tiny one-dimensional rips in spacetime known as strings" (Weinberg 1997, 20). On another thing-monist view, the wave function is the only thing that exists. Nevertheless I'll assume (D) at this point, for many philosophers believe it to be true. They also take it to give rise to a special and acute difficulty for any panpsychist theory: the so-called combination problem. So they might not be impressed by any argument for the primacy of panpsychism that assumed (B), thing monism, and so assumed that (D) was false.

3.10 The Hylal

With this in place, consider the proposal that

(17) experientiality is one possible fundamental kind of stuff and nonexperientiality is another.

This seems unexceptionable at first. 'Experiential' and 'nonexperiential' are mutually exclusive high-level type terms or 'kind-determinables' both of which, we may suppose, can have very different more determinate values.²⁶ We know this is so in the case of experiential stuff, in having sound experience, color-experience, taste-experience, and so on, and we naturally assume it may also be so in the case of nonexperiential stuff. We take it that there is wholly nonexperiential stuff in our universe and the supposition that there might be radically different kinds of wholly nonexperiential stuff (X-stuff, Y-stuff, Z-stuff) in other possible universes seems plainly coherent.²⁷

There is however an asymmetry when it comes to our understanding of the experiential and our understanding of the nonexperiential. In the case of experience we have a positive grasp of the sense in which all possible experience is, simply in being experience, the same fundamental kind of thing. Even if the particular qualitative character of Martian experience is radically unimaginable by us we still have a firm positive grasp of the fundamental kind of thing it is simply in knowing in general what experience is. When we consider the nonexperiential, by contrast, we suppose that we could possibly have a good grasp of the fundamental nature of the local nonexperiential stuff while really having no idea at all about the fundamental nature of X-stuff or Y-stuff or Z-stuff in other possible worlds. We know what experience is in an extremely general but still positively substantive way that allows us to see that experientiality constitutes a single fundamental kind, a single fundamental natural kind—a single fundamental qualitative kind, one might say, using 'qualitative' in a highly general way that has nothing specially to do with experience. The trouble is that we don't know what the nonexperiential is in the same sort of way, a way that allows us to say that it constitutes a single fundamental natural kind.

It's not hard to see why this is so: it's that we don't have a positive, substantive, general conception of the nonexperiential at all. 'Nonexperiential' is a merely negative, maximally general word that can as far as we know sweep up radical qualitative differences that don't fall under a single qualitative kind in the way that all kinds of experience knowably do.

How can we adjust the proposal so that it concerns two genuine fundamental kinds? It suffices to relativize the experiential/nonexperiential opposition to a particular universe, e.g. our own, replacing the maximally general negative term 'nonexperiential' by a more specific positive term that denotes the particular fundamental kind of nonexperiential stuff we take ourselves to encounter in our actual world.

Which term will suit? We obviously can't use the term 'physical' or 'material,' as real materialists who hold that experientiality is wholly physical. I propose 'hylal,' derived from the old Greek word for wood, which came to be used as a general term for matter conceived of as something entirely nonexperiential (consider Berkeley's 'Hylas'). All we need to stipulate for present purposes is that 'x is or has hylal being' entails 'x is or has nonexperiential being' whatever else is or isn't true of x.

3.11 Experiential-Hylal Monism?

With the term 'hylal' in place, (17) becomes

(18) experientiality is one possible fundamental kind of stuff and *hylality* is another.

We can then consider the proposal that

(19) reality may be *fundamentally* both experiential and hylal in nature

—where this is put forward as a stuff-monist proposal. On this view, the fundamental natural intrinsic properties of concrete reality include both experiential and hylal (hence nonexperiential) properties, even though (10) stuff monism is true. When we consider physical stuff, the only fundamental kind of stuff there is, we find both experiential stuff and hylal stuff.

If (10) stuff monism is true, as we are assuming, (18) rules out (19). For if (18) is true, (19) posits two fundamental kinds of stuff and is a version of dualism. So if one wants to continue to be a monist, and a real realist about experience, and hang on to nonexperiential stuff, in this case hylal stuff, as

many do (it's the only way to resist panpsychism), one has to suppose that the single fundamental kind of stuff may be fundamentally *both-experiential-and-nonexperiential* in nature: that experientiality and nonexperientiality, although essentially opposed, can possibly coexist as a single kind of stuff.²⁸

I'll call this position 'experiential-hylal monism'—'EH monism' for short: 'E' for experiential and 'H' for 'hylal.' Is EH monism possible? We can see straight away that no portion of E being can be H being, given that being H entails being non-E. And here I think we see the gain in clarity of giving up 'property' talk for 'being' talk.

Objection: It isn't a gain in clarity. It's an occlusion of a crucial metaphysical possibility. Of course H being can't be E being, but a portion of concrete being can possess both E *properties* and H *properties*. This is plain even when we restrict attention to 'fundamental natural intrinsic' properties, as you are doing here. Look, a *human being* can possess both E and H parts and properties.

Well, this could possibly be true of human beings—if there is some H stuff in our universe (so that pure panpsychism is false). But, first, it directly begs the question to say that it's obviously true because human beings certainly have both experiential properties and spatial properties, if one understands 'spatial' in the ordinary way according to which space-occupying properties are essentially or at least certainly nonexperiential properties. Our ignorance of the nature of the spatial rules out this proposal—even apart from the tendency among some leading cosmologists to question whether spacetime is fundamentally real.

Secondly, even if it could be true that things like human beings possess both E properties and H properties, I don't think it could be true in such a way that *EH monism* is or could be true. This, I think, is one of the places where property talk leads us grievously astray. It beguiles us into believing in the coherence of metaphysical possibilities that are in fact illusory, given *Sein ist Sosein*. It seems to allow for the possibility that a portion of concrete being may possess both fundamental natural intrinsic E properties and fundamental natural intrinsic H properties *without being ultimately wholly factorable into wholly E portions and wholly H portions*. In that case, however, some not-further-factorable or ultimate parts are both irreducibly E and irreducibly H, that is, impossibly, irreducibly both E stuff and non-E stuff.

Sein ist Sosein shows up the impossibility. Whatever one thinks about how properties may possibly coexist, concrete being is wholly qualitativity, concrete qualitativity, according to Sein ist Sosein, and E qualitativity can't be non-E qualitativity and conversely. So E qualitativity and H qualitativity can't possibly coexist in one nonfactorable portion of being (a portion of being that isn't ultimately made up of distinct nonoverlapping portions of E qualitativity and H qualitativity). In order to do this they would need something—the objectas-opposed-to-the-properties, the 'subject' or 'bearer' of the properties—that 'has' them and that is not itself wholly a matter of qualitativity. But there is no such thing—for *Sein ist Sosein*.

If this is right, EH monism fails. The attempt to describe it while respecting *Sein ist Sosein* pushes us inexorably back to dualism. (I trust that the point doesn't depend essentially on *Sein ist Sosein*)

The picture will continue to appeal—the picture according to which a portion of single-stuff being can have both E and H properties without being factorable into E portions and H portions. I think this is the way many aspiring real materialists tend to think—in a vague quasi-pictorial way about neural goings-on that are experiential goings-on. It's very easy to slip back into this, in my experience. One pictures the neural goings-on—the sweeping nets and waves of electrochemical activity flickering across great connected skeins of neurons—as having intrinsically irreducibly H (hence non-E) features. One then thinks that these intrinsically H goings-on are in at least some of their parts or features also E goings-on. But *Sein ist Sosein* blocks this when it's thought through. For again, and crudely, (i) things are in the end wholly 'made of' qualities (*Sein ist Sosein*), (ii) and E and H are incompatible qualities, so (iii) nothing can be made of both at exactly the same place.²⁹

There's wide scope for missing the point, given the plasticity of property talk. I can't hope to meet all objections or convince those committed to the traditional conception of properties. I think many who count themselves as materialists will be unable to give up the idea that we know in some fundamental—perhaps Moorean—way what space is. (I was unable to give it up in Strawson 2003a, §8.) We are confident we know in some deep way what space is, however wrong we also are about it, and in particular know that having spatial properties like shape properties essentially involves having non-E stuff properties, and so know—given that we accept that experiences are brain states—that there are things that have both E properties and non-E properties.

A first reply may be that a thing's particular shape isn't a matter of the intrinsic nature of the stuff it's made of, and that we already know that experiential stuff can be spatial stuff. The main reply targets the presumption that a thing can't possibly occupy space without having without some nonexperiential stuff being. This presumption may be deeply woven into our fundamental intuitive conception of space, as remarked, and some may be too deeply committed to it to take seriously the possibility that it may be false. They will have to face the fact that it appears to be incompatible with any realistic (genuinely experience-acknowledging) version of stuff monism that retains the idea that concrete reality is spatial.

3.12 The Untenability of Neutral Monism

EH monism isn't a version of neutral monism—the view, to quote Russell, that "both mind [E, on the present terms] and matter [H, on the present terms] are composed of a neutral-stuff which, in isolation, is neither mental [E] nor material [H]" (Russell 1921, 25). It's the precise opposite—not *neutral* monism but *doubly committed* monism, *both-and* monism as opposed to *neither-nor* neutral monism. What the two views have in common as monisms is that they want to accord the same reality status to E being and H being while remaining monist.

Can neutral monism do better than EH monism in this respect? Could E and H be genuinely real properties of things while somehow emerging from some more fundamental underlying stuff which is neither E nor H but rather let us say— ϵ (pronounced "ayn"), where to be fundamentally ϵ is to be fundamentally both wholly non-E and wholly non-H?

The short answer is no, but I'll spell it out a bit. ¿ can't be *neither E nor non-E*, on pain of logical impossibility. And it has to be non-E, since it would otherwise be E, and so not neutral between E and H (it would also be panpsychist). ¿ must therefore be a kind of non-E stuff which is different from H non-E stuff. There is no other possibility.

But this isn't a real possibility on the present view. It's ruled out by (7) No Jumps or equally (8) No Radical Emergence, which not only lead to

(9) E stuff can't emerge from wholly and utterly non-E stuff (e.g. ε)

but also to

(20) H stuff can't emerge from wholly and utterly non-H stuff (e.g. ε).

Objection: How can you rule out the possibility that something is in itself wholly non-E and wholly non-H but is nonetheless genuinely *proto-experiential* and *proto-hylal* in such a way that E and H can emerge from it? After all, you yourself allow there are deep respects in which we're radically ignorant of the fundamental nature of things.³⁰

I admit our ignorance but remain firm in commitment to No Jumps and No Radical Emergence. The idea that something can be wholly nonexperiential but nonetheless 'protoexperiential' will always seem attractive. It is after all the standard view of the evolution of consciousness like ours, according to which biological experientiality (human or canine or feline, etc.) evolved from wholly nonexperiential origins. But No Jumps and No Radical Emergence are part of the deep structure of naturalism. We have no reason to believe that nature ever makes ontological jumps of the sort forbidden by (7) and we have very good reason to believe that it doesn't. I don't, however, need to make this move. All I need to do is to reissue the polite inquiry I made earlier. Why does anyone think anything nonexperiential exists at all? I think the polite inquiry is devastating, and I'll return to it. I know that some will be unimpressed by it, and by the commitment to (7) and (8), so it's fortunate that there's another way of showing the inadequacy of neutral monism.

The term 'neutral monism' is used in many ways, most of which appear to be ultimately panpsychist or 'idealist'. But there's one central straightforwardly ontological way of understanding what it is (Russell and James seem driven principally by empiricist epistemological considerations) which appears to be ruled out by Sein ist Sosein. For according to genuinely ontological neutral monism, E and H are fully and unqualifiedly real, natural, categorical features of concrete reality-they're irreducibly real features, not just appearances of some sort, even though there's supposed to be a key sense in which they're not fundamental features. Now Sein ist Sosein states that the complete stuff being of a thing at any time isn't really distinct from the stuff being of its real, natural categorical propertiedness at that time. So, given Sein ist Sosein, it seems that we can't really defend any sense in which the fundamental nature of concrete reality is ultimately neither E nor H. The words 'ultimate', 'fundamental', and 'intrinsic' can't help (one can presumably add 'intrinsic' in the sense of 'non-relational' to 'real, natural, and categorical' above). Nor does one need to endorse fullblooded Sein ist Sosein to reach this result; it's enough to hold that the real, natural categorical propertiedness of x is at least part of what constitutes the being of *x*.

It seems, then that neutral monism can't help with the 'mind-body problem,' when it's understood in this natural, straightforward ontological way. So if EH monism is no better, as I have argued, it looks as if we must either go back to dualism, which is not I think a serious option, or head in the direction of panpsychism.³¹

3.13 Experience Entails an Experiencer

So here I stand—a naturalist materialist monist who's wondering about the nature of concrete reality and who knows that the only general thing he knows for certain about concrete reality is that experience exists. I find myself being pushed to acknowledge that panpsychism is the most plausible form of monism or indeed materialism. I'm aware that

(21) experience entails an experiencer

so I'm going to have to allow that there are as many experiencers as there are genuinely ontologically distinct portions of experience—even though this may appear to make things more difficult for me as a fledgling panpsychist.

Some philosophers have questioned (21)—wrongly because all experience is necessarily experience-*for*; experience for someone-or-something. Experience is necessarily experiencing. It's necessarily had, felt, experienced by something. In this immoveable sense there is necessarily an experiencer whenever there's experience. So anyone who prefers the term 'panexperientialism' to the term 'panpsychism,' on the ground that 'panexperientialism' allows for the possibility that there can be experience without an experiencer, has gone wrong (in a way that isn't endorsed by Hume, it should be said, or by Buddhists). Note that to insist that an experience entails an experiencer isn't to claim that the experiencer must be irreducibly ontically distinct from the experience or last longer than the experience. It's not to favor any particular hypothesis about the actual concrete realization of the *experiencer/experiential-content* structure that is attributable to any episode of experience.

3.14 A Global Replace

So here I am. I already know that the most parsimonious hypothesis compatible with the data is that concrete reality—the stuff that realizes the concretely existing structure that physics picks up on—is wholly a matter of experience, experiencing, experientiality. Experience like ours certainly exists and it follows, given No Jumps or No Radical Emergence, that experience must be among the fundamental properties of concrete reality. (To try to hold on to nonexperiential being by holding that reality is nonexperiential in its fundamental nature but is nevertheless and at the same time 'protoexperiential' seems to be to try to paper over a crack in reality with a word. The crack—or chasm—remains untouched.)

So when it comes to considering the question of the fundamental nature of concrete reality the choice lies between supposing that both experientiality and some form of nonexperientiality like hylality are among the fundamental properties and supposing that only experientiality is. I haven't been able to make sense of the dual option, compatibly with retaining monism, and I don't think there could ever be a good argument for dualism, so long as the two stuffs posited by dualism are supposed to interact causally (briefly, I don't see what argument could undermine the claim that causal interaction is a sufficient condition of same substancehood). So I seem to be forced into panpsychism.

Can this last position really be said to be a form of materialism? Surely—the point should be familiar by now. Many materialists hold that all concrete being

is simply energy existing in one form or another—that is, (1). The panpsychist proposal is simply that the intrinsic nature of this energy is experientiality. The panpsychist hypothesis performs a 'global replace' on the objects of physics as ordinarily conceived. In so doing leaves the whole of physics—everything that is true in physics—in place. So too for all the other sciences. I'm a robust realist about physical reality, the theory of evolution, and so on, but I know of no argument that gives us any good reason to suppose that there is any nonexperiential concrete reality.

The claim that experience is all that exists isn't the incoherent claim that everything that exists exists only in or 'in' some mind or other (that's incoherent because a mind can't exist only in or 'in' itself). It has nothing to do with Berkeleian idealism, or phenomenalism, and it certainly isn't committed to the implausible view that tables and chairs are subjects of experience. It leaves the physical world untouched, as 'out there,' relative to each one of us, as it ever was—however inadequate our idea of its *Existenzraum* or dimensionality.

Objection: so there's no distinction between materialism and what amounts to a form of 'absolute idealism.'

There is a distinction if 'absolute idealism' implies (B) thing monism; but not if it's simply a form of pure panpsychism. I hope you don't think this is comic or absurd, because it looks as if it's materialism's best guess as to the nature of the concrete reality about which physics says many true things. Eddington and Whitehead saw this clearly nearly 100 years ago. You don't have to call it 'materialism' ('physicalism') if you don't want to. I continue to call it 'materialism' ('physicalism') because, once again, concrete reality understood in this way is what physics describes in its own magnificent and highly abstract way and says many true things about ($e = mc^2$, the inverse square laws, the periodic table, etc.), things which I take to hold good of everything that concretely exists.

Objection: But still—why not suppose that the basic nature of concrete reality is *non*experiential rather than experiential?

In that case we face again all the problems posed by No Jumps and No Radical Emergence. Suppose those problems solved. Then I reply to your question—'Why suppose that the basic nature of concrete reality is experiential?'—with another question: 'Why suppose that it's nonexperiential—either in its basic nature or in any respect at all?' What evidence is there for the existence of nonexperiential reality, as opposed to experiential reality? None. There is zero observational evidence for the existence of nonexperiential reality may that each of us encounters a great deal in concrete reality that is not his or her own experience. Nor will there ever be any. All there is, is one great big wholly ungrounded wholly question-begging theoretical intuition or conviction.

Objection: There isn't any evidence that the intrinsic nature of reality is wholly experiential either.

True—but we know that some of it is experiential. We know it for certain because

(22) In the case of experience, the having is the knowing.

To have experience is not only to be directly acquainted with the fundamental nature of experience—at least in certain respects. It's also of course to know that the experiential exists. The view that there is any nonexperiential concrete reality is, by contrast, wholly ungrounded. It's a radically and irredeemably verification-transcendent belief. Hume knew this. So did many others, including Quine, who famously judged that physical objects that are assumed to be nonexperiential are "posits comparable, epistemologically, to the gods of Homer" (Quine 1951, 44)

3.15 Ignorance and Repugnance

Objection: It's an old point that there's an evident and fundamental 'repugnance' or incompatibility between the spatial on the one hand and the conscious or experiential on the other. We encounter things in space, we know them to be in space, and since we have powerful reasons, given the repugnance, for thinking that spatial things can't in themselves be experiential things, or at least can't be wholly experiential things, we have decisive evidence that there is nonexperiential reality.

This issue arose in §10. One doesn't have to agree with Kant that space isn't ultimately real, but just a 'form of sensible intuition,' to grant that we may be very ignorant of the nature of space or spacetime. The intuition of knowable repugnance went south long ago. It was understandable in the seventeenth century, the age of classical contact mechanics, but it doesn't look very good in the twenty-first. We know the experiential is real and we also know—about as well as we know anything in science—that it's literally located in the brain:

(23) human experience is neural activity.

This is by now far beyond reasonable doubt. So we know, about as well as we know anything in science, that the spatial can be experiential—given that the world is spatial. And in the present state of our knowledge we have to treat 'space' and 'spacetime' as names for some real dimensionality whose nature

we aren't clear about, although we know that it must be such as to allow the existence of experientiality.

What can we say in general about this dimensionality, given that pure panpsychism is the most parsimonious hypothesis about the nature of concrete reality? Not much but not nothing. As pure panpsychists we may take it that (a) the dimensionality of the concrete real, however ill understood by us, is something that fits smoothly with (b) the nature of the concrete real conceived of as nothing but experientiality in exactly the same way as the way in which (c) the dimensionality of the concrete real understood as spatial in the conventional way is seen to fit smoothly with (d) the nature of the concrete real understood as good old fashioned nonexperientially propertied extended physical stuff (plainly any difficulty lies in the idea of space, not of time).

It may be said that we must retain the idea of dimensional *position*, even when we figure the dimensionality of the concrete real as something that fits smoothly with the idea that the concrete real is wholly experiential, because the idea of position and difference of position is essentially built into the idea of dimensionality. It may then be said that the property of having some dimensional position, at least, is something essentially nonexperiential. But this may be readily granted because it raises no difficulty for the idea that the whole intrinsic nature—stuff nature—of the concrete real is a matter of experientiality.

We find it quite incredibly hard to think clearly about these things, as Russell stressed. Almost all of us are in his words "guilty, unconsciously and in spite of explicit disavowals, of a confusion in (our) imaginative picture of reality" (Russell 1927, 382). Even when we admit and dwell on our ignorance—even perhaps, when we have seen the force of the argument that all that concretely exists is the wave function—we tend to revert to a conviction that we have a basic grasp on things that allows us to be sure that the matter/energy whose spatiotemporal manifestations are all around us couldn't literally be nothing but experientiality.

3.16 Pictures of Matter

There's no direct remedy for this. But there are some mental exercises one can perform. It's helpful to keep a few well-known physical facts vividly in mind and constantly remind oneself of them when facing the 'mind-body problem.' Consider first the fact that the spatial volume occupied by one's brain—equivalent to the volume of a sphere about five inches across—is, intuitively, almost completely empty. (More accurately, it's almost all 'quantum vacuum' and is arguably a *plenum*, like all spacetime, that is, the precise opposite of a vacuum; while still being, intuitively, almost completely empty.) Add the fact that it contains about 100 billion nonneuronal cells, although it's almost completely empty, and an approximately equal number of neuronal cells that have up to a thousand trillion synaptic connections between them—plus the fact that about 700 billion solar neutrinos (and heaven knows what else) pass through it every second.

From one intuitively natural perspective matter is quite astoundingly insubstantial, an intricately shimmering almost-nothing. And this is so even when we consider a pebble or a mountain. When we go on to consider a brain we find many further layers of staggeringly intricate organization—in an almost entirely empty space. Such is matter. Such is the material brain. It helps to maintain this picture when we're wondering how experience can be physical. It helps to resist the picture of a mammalian brain as a 'sludgy mass,' a piece of meat that can be diced and fried with garlic; although it's also that.

I think it's also very important to habituate to *Wesen ist Werden*—the processual view of reality. It has to become something more than book learning. The same goes for *Sein ist Sosein*. We need to be able to put aside as far as possible the object-property distinction that serves so well in many other areas of philosophy and everyday life but easily leads us into an intractably misleading picture according to which, when it comes to the 'mind-body problem,' we have to think first that we have a *thing*, a physical thing, and then have to wonder how such a thing can possibly have experiential *properties*.

The improved picture represents matter soberly and realistically as an almost inconceivably sparse shimmering skein of energy, energy-stuff. On one view, this is what spacetime is. But insofar as this picture has positive imaginative (quasi-pictorial) content for us, it still builds in some version of our ordinary conception or picture of space. And if we now try to exert ourselves further imaginatively, in the way philosophy so regularly demands, in the attempt to put aside any standard conception or picture of space, we're returned to the 'epistemic structural realist' point that we know nothing at all about the intrinsic nature of the physical in so far as as its intrinsic nature is more than its scientifically detectable structure; except of course, and again, and as always, when we have experience. As Russell says: "we know nothing about the intrinsic quality of physical events except when these are mental events that we directly experience" (Russell 1956, 153).

We have then to allow again that our picture of space may be profoundly misleading. This doesn't prevent us from appreciating the extraordinary insubstantiality of the physical, even if we can't shake off the conventional spatial imaginings that come with it. But those who (like my former self) can't shake their commitment to the idea that we know what space is in some truly fundamental and Moorean respect may simply be unable to engage fully with the 'mind-body problem.'

3.17 Fungibility

We know experience exists. We've assumed that stuff monism is true and that everything is physical. It seems that

(24) there is no good reason to believe that anything nonexperiential exists³²

because there's zero evidence for the existence of nonexperiential reality. One thing we now need to consider is the idea that

(25) all physical stuff is *fungible*

in the sense that any form of it can in principle be transformed into any other—so that if for example one broke hydrogen down into leptons and quarks one could reassemble it as gold.³³ If this is so then it seems plausible to suppose that all physical stuff can potentially be part of what constitutes is—experientiality like ours in living conscious brains like ours, that is, that

(26) all physical stuff can constitute (be) experientiality, experiential being.

And if so, then—given (8) that there is no radical emergence, given that one can't get the experiential out of the nonexperiential by any kind of rearrangement of the nonexperiential—it seems we can advance from (24), according to which there is no good reason to believe that anything nonexperiential exists, to

(27) we have good reason to believe that nothing nonexperiential exists.

But never mind (27), which does after all rely on (8). The weaker (24) is enough for now. One of the most important experiences that a philosopher brought up in the (recent) Western tradition can undergo is the realization that (24) is true: the belief in irreducibly nonexperiential reality has no respectable foundation, even given a fully realist commitment to belief in an external world of tables and chairs—a world that exists wholly independently of one's own mind and one's experiences—and a conviction that physics and cosmology and indeed the other sciences—get a very great deal right about the nature and structure of reality. The experience is life-changing, philosophically. It's natural for many to think that it is nevertheless essentially theoretically cheaper to suppose that the fundamental nature of concrete reality is nonexperiential—rather than splurging on universal experientiality. But this is simply a mistake. The postulation of fundamental nonexperientiality not only commits one to something for whose existence there is no evidence; it also commits one to belief in radical emergence. It's far more expensive, theoretically speaking.

3.18 No Mystery

Many say that experience (consciousness) is a mystery. But what is mysterious? We know what experience is. We know exactly what certain types of experiences are simply in having them. More precisely: we know exactly what certain types of experiences are considered specifically in respect of what they're like for us experientially. And, again, we not only know in this way what particular types of experience are. We also know what experience is generally considered—even though we have direct experience only of certain limited kinds of experience (see Sprigge 1999).

'We know exactly what certain types of experiences are considered specifically in respect of what they're like for us experientially.' The 'considered as' qualification makes room for the idea that our experiences have some further intrinsic nature that transcends what we know in knowing their phenomenological character in having them. It seems wise—necessary—to allow for this, given that experiences are neural goings-on, and given all the wonderfully precise numerical things physics and neurophysiology can say about them considered as things whose existence involves subatomic particles, atoms, molecules, individual cells, and so on. We needn't however suppose that any aspect of the being of our experiences that transcends what we know of their being simply in having them involves anything nonexperiential. The (pure) panpsychist proposal is precisely that all the subatomic, atomic, molecular and cellular energy phenomena about which physics and neurophysiology say true and extraordinarily precise numerical things are themselves experiential phenomena microexperiential phenomena.

Call these microexperiential phenomena *Es. Es* may have phenomenological features of which we have no knowledge, in having the kinds of experiences we have, even though they somehow conspire to constitute our experiences. The energy that is an electron is wholly a matter of experiencing, on the present view, but the specific phenomenological character of this experiencing may be radically unimaginable by us. This doesn't change the fact that we know its
nature in a fundamental general way. We do, because we know what experience is in a fundamental general way simply in having experience. The psychophysics (to give this term a new use—the point is that physics is psychics) of the universe is mysterious to us; we don't know how it is that energy is experientiality. But in God's physics asking how it is that energy is experientiality may be like asking how it is that energy is energy; and there is, for all our ignorance, a fundamental sense in which we know the nature of the stuff out of which everything is made in knowing what experientiality is.³⁴

I don't know anything about the Laws of Experiential Combination that govern the way Es constitute macroexperiential phenomena like human and canine experience. On one view Es undergo radical fusion in such a way that there's nothing more to their experiential being-hence their being tout court—than what we experience in having experiences. I don't see how this can be so, because phenomenological being is all there is to the being of experiences, according to pure panpsychism, and the complexity revealed by physics and neurophysiology isn't phenomenologically given in our experience. On another fusion-like view the fundents (the fundentia, the elements that fuse) somehow continue to possess some intrinsic experiential character of their own even as they unite in such a way as to jointly constitute experience like ours. One hypothesis is that these fusions or unities are what show up as gusts—waves, bursts—of synchronized activity in the brain. Perhaps they involve massive quantum entanglement effects or other strongly unificatory phenomena that can be identified as such by physics even though physics can characterize them only in nonexperiential terms. I don't know. What I do know is that we can't demand more intelligibility from the Laws of Experiential Combination (experiential chemistry) than we demand from quantum mechanics and physics in general.

What remains is the respect in which there is no fundamental mystery if panpsychism is true—even though we have no idea how the macroexperiential arises out of the microexperiential. Radical mystery is introduced only by the hypothesis that the intrinsic nonstructural or stuff nature of matter is (i) *non*experiential (hylal), at least in part, and hence radically distinct from anything we know in knowing what experience is, and is furthermore (ii) of such a nature that we are utterly unable to see how how it relates ontologically to experience (the 'explanatory gap').³⁵ We have, again, no idea of the intrinsic nonstructural stuff nature of the physical insofar as the physical is something that exemplifies the structures that physics detects.³⁶ The point is not just that the numbers and equations of physics don't capture the *whole* basic or essential nature of reality; it's that they tell us nothing about the nature of concrete reality insofar as its nature is more than its structure.

The fact that physics is full of mystery—things we can't claim to understand at all—is universally conceded quite independently of this point. Bohr, Einstein, Feynman, Penrose, Schrödinger, Wheeler all agree. No doubt the experiential is a mystery relative to physics. But to be a mystery relative to physics is to be a mystery relative to a mystery; and if something is a mystery relative to a mystery it need not itself be a mystery. It may be that it's only relative to a mystery that it looks like a mystery. And the point that physics is silent about the intrinsic nature of the physical, insofar as the intrinsic nature of the physical is more than its structure, is entirely general. It has nothing specially to do with experience. It holds equally on the supposition that the intrinsic (nonstructural) nature of the physical is wholly nonexperiential. So there's no special puzzle or problem in the fact that physics finds no place for experience (consciousness). It finds no place for any positive characterization of the intrinsic nonstructural features of concrete reality. "If you want a concrete definition of matter it is no use looking to physics" (Eddington 1928, 95).

3.19 Occam

So what should we real materialists do—if and when we try to do metaphysics? Like many I think we should start from something we know to exist and whose nature we know—the human experiential. It's beyond reasonable doubt that human experience is wholly a matter of neural goings-on and it seems no less clear that the most parsimonious scientific hypothesis about the nature of physical reality is that everything is experiential.

The experiential starting point isn't chosen for reasons of epistemological or ontological caution. It's just that it's usually best to start from something one knows to exist if one wants to try to give an account of how things are. I'm not particularly cautious when it comes to metaphysics. I'd postulate nonexperiential reality in a flash if I could see how postulating it could help in any way with any problem in real metaphysics or make a contribution to any view of how things are that we have any good reason to believe. As remarked, it's scientific orthodoxy that concrete reality consists entirely of energy. The present proposal, once again, is simply that the intrinsic nature of the energy is experientiality. I'm ready to change my mind if someone can show that the hypothesis that the energy phenomena that physics studies have some intrinsic nonstructural nonexperiential nature is superior to the hypothesis that their intrinsic nonstructural nature is wholly experiential—e.g. by showing some special difficulty in the hypothesis that they are wholly experiential. As far as I can see, however, there isn't a scintilla of a reason for postulating anything nonexperiential. Occam's razor, according to which one shouldn't as a theorist posit more entities than one needs to explain the data,

(28) entia non sunt multiplicanda praeter necessitatem

slices away the nonexperiential even if one is an all-out external-world realist.

3.20 The Combination Problem

This is naturalist monist 'global replace' materialist panpsychism. It denies the existence of something for which there is no evidence, that is, nonexperiential reality (eliminating the need to postulate radical emergence), even as it admits to being clueless about how biological (e.g. mammalian) experientiality emerges (nonradically) from the overall energy-experientiality of the fundamental constituents of brains.

At this point real-materialist opponents of panpsychism bring up the socalled *combination problem*.³⁷ They accept (10) monism, and (11)/(12) real realism about experience. They believe (13)/(14) that the equations and numbers of physics cotton on accurately to something real. And they endorse (D) 'smallism.' They also grant—indeed insist—that (21) every distinct experience necessarily involves a subject.

It's the last two claims that directly underlie the combination problem. The central idea is that a group of distinct experiencings or patches of experientiality, each of which necessarily has its own subject, can't possibly interact or fuse or coresonate in such a way as to constitute or generate a single experience with a single subject. Why not? Simply because

(E) a plurality of subjects can't possibly combine to form or generate a single subject.

If (E) could be proved true, I'd give up (D), which is highly questionable, according to certain leading conceptions of physics and cosmology. But as things stand I see no good reason to accept (E) even given (D). When I figure matter as a shimmering skein of energy/experientiality as best as I can; when I factor in my imperfect lay grasp of the phenomena of quantum entanglement and the extraordinary difficulties that arise when it comes to questions of synchronic and diachronic identity in fundamental physics; when I factor in my imperfect grasp of the fundamentality of field-theoretic conceptions of those phenomena that lead us to talk in an arguably misleading way of 'particles'; when I form an imaginative picture of small patches of local influence fusing into larger transient local patches (adding, perhaps, a flavoring of dark energy and dark matter)—I can't feel any deep difficulty in the subject combination problem.³⁸ Once again, we can't expect the Laws of Experiential Combination to be more open to human understanding than the laws of quantum mechanics.

Nor can I feel any deep difficulty in the 'grain problem'—the fact that "there seems to be a profound *structural mismatch* between the contents of one's consciousness at any given time, and what science would tell us is simultaneously going on in the brain" (Lockwood 2003, 453).³⁹ So too, when I consider the two groups of three parameters that account for all the colors and sounds, or the five that account for all the tastes, or the combinatorial possibilities of leptons and quarks—the astonishing variety of stuffs (lead, neurons, marshmallow) they constitute—I feel no difficulty in what Chalmers has called the 'palette problem' (see Chalmers this volume): the numerical gap between the relatively small number of fundamental entities postulated in the standard model of physics and the seemingly vast number of different types of experiences.⁴⁰

The basic point is simple. (1) We have no good grounds for thinking that we know enough about the physical to have good reason to think that these problems are serious difficulties for panpsychism.⁴¹ Kant makes the point well: if someone rejects materialism (or equally micropsychism) and argues for a simple immaterial soul "merely on the ground that the unity of apperception in thought does not allow of its being explained [as arising] out of the composite, instead of admitting, as he ought to do that he is unable to explain the possibility of a thinking nature (einer denkender Natur)," why should not the materialist [or equally the micropsychist], though he can as little appeal to experience in support of his possibilities, be justified in being equally daring, and in using his principle to establish the opposite conclusion?" (see Kant 1787, B417–18, my emphasis).⁴²

(2) More positively: we have strikingly good grounds for thinking that many of our intuitions of irreducible ontological separateness and distinctness are profoundly mistaken. (3) The problems that arise for a physicalism that postulates fundamental nonexperiential reality and so rejects panpsychism are far greater than the problems that arise for panpsychism (e.g. zero evidence for nonexperiential reality, contravention of No Jumps and No Radical Emergence).

I also believe (with Descartes, pre-Critical Kant, and many others, including William James) that there's a metaphysically primordial way of thinking about what a subject of experience is given which there is, in the case of any particular episode of experiencing, no real distinction between the subject of experience or experiencer and the experience or experiencing (see e.g. Strawson 2003b). This may contribute to my failure to feel worried by the combination

problem. I don't, however, think that this particular belief is indispensable to the lack of worry—except insofar as it's linked to the *Sein ist Sosein* claim.

And now a further question arises. The idea that there's nothing but experiential reality is supposed to give rise to certain distinctive problems; but how can the supposition that there is nonexperiential reality improve things?⁴³ It may be said that it does remove the supposed 'combination problem.' But I've already lost any sense that we have good reason to think that this is a serious difficulty, and the problem of how pluralities of distinct *non*experiential processes can combine to form necessarily single-subject-involving experiencings like your and my current experience looms no less large, given that it requires (among other things) radical emergence of the experiential from the nonexperiential.

3.21 The Primacy of Panpsychism

I'm not claiming to know that there is no nonexperiential reality. I'm just considering the most plausible scientific hypothesis—'global replace' real materialist panpsychism—and wondering why the self-styled hard-nosed naturalists of our day (a) deny the existence of something that knowably exists and (b) assume the existence of something for which they have no evidence: nonexperiential reality. My bet is that

(29) everything is experiential

—that the intrinsic (nonstructural) nature of the energy that is widely agreed to wholly constitute physical reality is experientiality.

I can't prove this, of course. Some will think that the combination of (7) No Jumps and (28) Occam's razor is very close to proof but I'm content to argue for something weaker—for

(30) the primacy of panpsychism

as advertised at the outset-the view that

(30) we should favor panpsychism over all other substantive theories of the fundamental nature of reality.

It's not the only game in town, when it comes to speculating about the ultimate nature of reality—unless William of Occam is the sheriff—but it's the best theory we have.

3.22 Awareness of Awareness; the World-Knot

I'll end with a very brief, more positive, and wildly speculative thought. It begins with Aristotle, at least in the Western tradition, who observes that "if we are aware, we are aware that we are aware" (*Nicomachean Ethics* 9.9.1170a29b1).⁴⁴ I'm going to take this claim to be correct in saying that all experience, all *awareness*, as I'll now also call it—using 'awareness' to refer only to conscious awareness and taking it to be synonymous with 'experience' used as above— somehow or other involves awareness of that very awareness.

This can be read in at least two ways, as is well known: in a higher-order way and a same-order 'self-intimationist' way. I favor the same-order view, the view that

(31) all awareness on the part of a subject comports awareness, on the part of that subject, of that very awareness

where 'comports' is used to mean something like (and at least) 'contains within itself' (as in French *comporter*); so that the awareness of awareness isn't any-thing ontically over and above the awareness considered as a whole.

We can rewrite (31) more simply as

(31) all awareness comports awareness of that very awareness.

I'll call this the Self-Intimation thesis.

There's an enormous quantity of discussion of this matter. I'm not going to add to it here.⁴⁵ I'm simply going to endorse the Self-Intimation thesis in order to propose a further Very Large Step: perhaps the self-intimation—the fundamental self-reflexivity, the *Fürsichsein*—characteristic of experience is of the essence not only of experience—mind, consciousness—but of all concrete being. Perhaps it's only this kind of turnedness-on-itself that can catapult or bootstrap being into being. This would explain why *Sein*—*Ansichsein*—is *Fürsichsein*: there is no other possibility:

(32) all concrete reality is necessarily experiential.

Catapulting and bootstrapping are bad metaphors insofar as they suggest that experience is somehow *causa sui*—the cause of itself. I don't think anything can be *causa sui*. A thing can be somehow self-sustaining, perhaps, but not self-caused. Slightly better, perhaps, is the proposal that this sort of selfreflexivity or self-relatedness or self-intimation is an internal sprungness—a self-sprungness—that characterizes not only awareness but being in general. It's what holds off its collapse into nothing (as it were). It is perhaps one good way to characterize what energy is, and the whole of concrete being is energy in one form or another: *Stoff ist Kraft*.

This perhaps is the real 'world-knot'—Schopenhauer's term for the point of contact between subject and world. It really is a kind of knot inasmuch as a knot is essentially turned on itself. This perhaps is the real remarkableness of experience as it emerges in our attempt to theorize about it. And it's nothing other than the remarkableness of concrete being.

This is entirely speculative. The notion of being self-sprung is metaphorical. But I think that something about it smells right-the idea that the 'selfsprungness' or 'self-intimation' of experience is the fundamental form or selfsustaining structure of the energy which is concrete reality. Self-sprungness makes—constitutes—force, and Stoff ist Kraft. Matter—more generally, the physical, all concrete being—is force or activity or power or energy. Matterforce is essentially dynamic, being is essentially becoming: Wesen ist Werden. We travel smoothly down the chain of terms which—it now appears—forms a circle: a panpsychist circle. We already know that we neither have nor can have any good reason to think that anything nonexperiential exists in concrete reality and we've now noted a fundamental feature of experience—a kind of selfloopedness that seems uniquely characteristic of experience—that offers itself as a fundamental feature of any kind of concrete being at all. With Eddington, Russell, Whitehead, and many others, including Spinoza and Leibniz, and perhaps Kant, and many others, I suspect we're wrong to think that awareness or *Fürsichsein* is a special—rare—feature of the universe. On the present view it's the most common thing there is. In fact it's the only kind of thing there is. All being in-itself, that is, all being, all being period, is being for-itself. This is an essential part of its intrinsic or ultimate nature. This is what energy is, the energy treated of in physics, the energy of which matter is one form among others, and about whose intrinsic nature, over and above its structural nature, physics has, provably and forever, nothing to say.

"Apart from the experiences of subjects there is nothing, nothing, nothing, bare nothingness" (Whitehead 1929, 167). I don't know exactly why Whitehead came to this conclusion, but we know—to say it one more time—that experience exists in the universe, and we don't know that anything else exists. This isn't any sort of argument that nothing nonexperiential exists, but all those who are genuinely committed to monism ought to prefer the hypothesis that everything is experiential to all hypotheses that suppose that the fundamental nature of reality is wholly nonexperiential because all these hypotheses require that one posit radical emergence. They require experiential phenomena to emerge from phenomena that are in themselves wholly and utterly nonexperiential. Long familiarity with a picture according to which experientiality emerged from nonexperientiality in the course of biological evolution has softened our thinking in such a way that we can no longer clearly see what an extravagant hypothesis this is—especially for someone who is convinced, as I am, of the truth of the theory of evolution.⁴⁶

Notes

- This paper is a composite of talks first given at the Towards a Science of Consciousness conference in Tucson in 2010 and a conference in München in 2011, and subsequently at Wollongong University (Australasian Association of Philosophy conference), New York University, Rice University, Rutgers University, the University of Mississippi, and the Czech Academy of Sciences in Prague. I am grateful to many for their comments. I remember in particular Pierfrancesco Basile, Dave Chalmers, Sam Coleman, Philip Goff, Keith Turausky, Hedda Hassel Mørch, James Ladyman, Pat Lewtas, Anna Marmodoro, Michelle Montague, and Udo Thiel.
- 2. In this paper I'm concerned with concrete being—the universe. I don't know whether it's helpful to say that there is abstract being, as opposed to and in addition to concrete being, but I'm going to put this question aside.
- 3. See also, strikingly, Faraday (1844, 140ff), Bohm (1957, §1.6), and many others (when I cite a work, I give the date of first publication, or occasionally the date of composition, while the page reference is to the edition listed in the bibliography). I'm inclined to include Plato, who holds that "being is nothing other than *dunamis*" (Plato *c* 360 BCE, 247d-e), that is, potency, power, force. But this would need to be vigorously argued, given the way in which Plato distinguishes between *dunamis*, potency, and *energeia*, actuality.
- 4. Kant 1781–87: B288. Kant's formulation is entirely general and can be taken nontemporally. Note that I use 'phenomenon' in a standard way as a completely general word for any sort of existent, and entirely put aside its traditional meaning of *appearance* (I have learnt that this can cause great confusion).
- 5. The existence of power can't be equated with the existence of something 'merely' dispositional—given one common understanding of the word 'disposition.' Just as there's no energy without power so there's no power without actual, live energy (it would be superficial to think that the existence of *vis inertiae* doesn't involve the existence of energy).
- 6. I'll use letters rather than numbers for primary propositions that I don't positively endorse in this chapter although I think some of them may be true.
- 7. Samuel Alexander endorses (A): "Space-Time is the stuff of which matter and all things are specifications" (Alexander 1924, vi). I'm leaving aside the 'relational' conception of space because I don't think anyone has ever managed to make sense of it as a metaphysical position.
- 8. See Strawson *forthcoming b*.
- 9. "Wirklich ist, was wirkt, was eine Macht, eine Potenz ist" (Frauenstädt 1840, 341).
- Schopenhauer 1819–59, \$1.4; matter is "causality itself, objectively conceived" (Schopenhauer 1819–59, \$2.1.4); "matter is throughout pure Causality, its essence is Action in general" (Schopenhauer 1813, 97).
- 11. Two quick points on this view. (1) You can't vary the nomic circumstances of a thing x the laws of nature governing x—while keeping x's nature constant, because the laws are essentially constitutive of its nature. (2) Even if you could, there would be no independence of behavior from intrinsic nature. For x would behave in the way it did, say W1, in nomic circumstances N1, wholly because of its intrinsic nature, and it would behave in way W2 in nomic circumstances N2 wholly because of its intrinsic nature; etc.
- See Strawson 2008, 279–81. We lose hold of the key point if we take the identity claim in (3*) to be just a version of the 'bundle' theory of objects. The bundle theory of objects as

standardly presented will always seem intuitively unacceptable, and rightly so, because it retains the everyday conception of, and distinction between, 'object' and 'property'—the very conception that is undermined by insight into the identity claim.

- 13. There is a basic sense in which all consciousness is a form of self-consciousness, and *Fürsichsein* may also be linked more specifically to the notion of self-consciousness.
- 14. One may then suppose that mind—mind proper—is always the result of some sort of evolutionary process, although experience is not.
- 15. On this qualitative construal (7) is wholly compatible with all quantum phenomena and all phenomena cited in support of 'saltationism' in the theory of evolution.
- 16. For some arguments see Strawson (2006, 60–67); see also Seager (2012), Chalmers (this volume). Compare Jackson's arguments for *'a priori* physicalism' (cf. e.g. Jackson 2003).
- 17. It is arguable that Descartes holds (B) with respect to concrete material reality; Spinoza holds it with respect to all concrete reality. Among those who endorse this view today are Horgan and Potrč (2008). Schaffer calls this view 'existence monism' (see e.g. Schaffer 2007; 2010).
- 18. I'm putting aside ethics where I'm not a naturalist—*if* being a moral realist excludes being a naturalist.
- 19. One can always substitute the word 'consciousness' if one wishes.
- 20. See also Greg Rosenberg (1999; 2004) and Ladyman et al. (2007). Compare Descartes: "all the properties" of material things "which I clearly and distinctly understand are . . . comprised within the subject matter of pure mathematics" (Descartes 1641, 2.55). Poincaré puts the point very vividly in chapter 10 of *Science and Hypothesis* (see Poincaré 1903).
- All theories that claim to give a reductive account of experience in terms of nonexperiential phenomena—for example, behaviorism in all its forms and all full-on versions of functionalism—deny the existence of experience. They claim not to on the ground that reduction is not elimination; but reduction is elimination in this case (see Strawson *forthcoming a*).
- 22. See e.g. Ladyman et al. (2007).
- Ladyman and Ross appear to bite this bullet: "we reject the dichotomy between the abstract and the concrete, and between the substantival and the structural" (Ladyman et al. 2007, 186).
- 24. It's also arguable that (C) entails (B) on the ground that the universe must be correctly describable as a single structure if (C) is true.
- 25. See e.g. Wilson (2004), Coleman (2006).
- 26. Compare 'color,' 'shape,' and 'animal,' each of which have many more determinate 'values'—'red,' 'blue,' 'round,' 'square,' 'cat,' 'dog.'
- 27. It may be, in fact, that nonexperiential concrete stuff is not possible. I'll consider this suggestion at the end.
- 28. Compare Regius's suggestion (there is little doubt that he is reporting a view that Descartes also entertained): 'some philosophers ... hold that mentality/consciousness [cogitatio] and extension are attributes which are present in certain substances, as in subjects; [and] since these attributes are not opposites but [merely] different, there is no reason why mentality/consciousness should not be an attribute of some sort co-existing with extension in the same subject, though the one [attribute] is not included in the concept of the other. For whatever we can conceive of can exist. Now, it is conceivable that mentality is something of this sort; for it does not imply a contradiction. Therefore it is possible that the mentality is something of this sort. So those who assert that we clearly and distinctly conceive human mentality as necessarily really distinct [in Descartes's sense] from body are mistaken' (Regius 1647, 294–95).
- 29. *Objection:* we allow that monism—physicalism—may be true even if there are irreducibly different (perhaps essentially noninterconvertible) 'fundamental particles.' Why can't we similarly allow that monism may be true when there are irreducibly different (essentially noninterconvertible) types of fundamental stuff—E stuff and H (non-E) stuff?

The first reply is a question: why bother, given that there is no good theoretical reason to posit H stuff? A further reply is that particles are emergent phenomena according to quantum field theory, plausibly all 'made of' the same kind of stuff. One can also question the noninterconvertibility of fundamental particles (see the discussion of 'fungibility' in §16)

and note that the view that the fundamental entities are strings with different vibrational characteristics creates no evident difficulty for stuff monism. More generally, we take ourselves to have strong reasons for holding that all the fundamental particles are of the same fundamental kind. In the case of E stuff and H stuff, by contrast, we know the fundamental nature of E stuff, and H stuff is *defined* as non-E.

- 30. See Stoljar 2006a. Stoljar points out that I sometimes appeal to radical ignorance in argument, e.g. citing 'the silence of physics' against the view that we have any reason to believe in nonexperiential being, and at other times reject appeals to radical ignorance in arguments made against me, e.g. when standing up for No Radical Emergence. This is true but it is not I think a difficulty.
- 31. In Strawson (2003a, 50) I argue that we can never have good reason to prefer dualism (or any pluralism) over monism, so long as we posit causal interaction between the two supposedly distinct substances. Note that *Sein ist Sosein* also entails the incoherence of so-called property dualism.
- 32. This is not the unintelligibility claim made by Berkeley (1710; see Foster 1982 and Robinson 2009), according to which the notion of the nonexperiential is wholly unintelligible: I'm happy to allow that the general notion of nonexperiential reality is wholly intelligible.
- 33. I don't know if this is unrestrictedly true; one contrary line of thought runs as follows. It seems plain that (1) experientiality has something essentially to do with electricality in our world at least, and perhaps in all possible worlds. One might accordingly suppose that (2) all electricality is experientiality, and perhaps also that (3) all experientiality is electricality. One might then hypothesize that (4) ultimate constituents of reality lacking electricality (chargeless particles—neutrinos, photons, 'chameleons') are intrinsically nonexperiential and can't directly constitute (be) experientiality, and further (5) that fungibility might fail between ultimate constituents of reality possessing electricality and ultimate constituents lacking it.
- 34. It may be doubted whether there is any robust sense in which an electron is a genuinely individual or persisting thing. On one reading, relativistic quantum field theory has it that the phenomena that lead us to talk in terms of particles are simply manifestations of the quantization of the energy of fields, and aren't well thought of as entities that can be said to endure over any significant period of time.
- 35. See Leibniz 1704, §17; Levine 1983.
- 36. Russell writes that "we know nothing about the intrinsic quality of physical events except when these are mental events that we directly experience" (Russell 1956, 153). Unlike Russell, I take a concrete thing's structural nature to be part of its intrinsic nature.
- The combination problem was clearly stated by Lucretius 2,000 years ago (Lucretius 50, 2.865–990). See also Collins and Clarke (1707–1718), James (1890, ch. 6), Goff (2006, this volume), Chalmers (this volume), Coleman (this volume).
- 38. See Seager's discussion of 'combinatorial infusion' (Seager 2010 and Seager this volume).
- 39. See also Lockwood 1993. For one interesting form of reply to the 'grain problem,' cf. Stoljar (2006b, 117–18).
- 40. I'm mindful, also, of Turausky's suggestion that particular experiences may be formed by subtraction—reduction—sculpting—of a base of experiential 'white noise' (see Turausky *unpublished*).
- 41. Remember that these proponents of the combination problem hold that the experiential is wholly physical—or at least, as monists, that it is wholly of the same stuff as whatever stuff they postulate.
- 42. Kant undermines the view that we can know that the mind or soul or thinking subject is a single substance in the Second Paralogism. He grants—stresses—the sense in which the thinking subject is something that is necessarily single in the activity of thought or experience, and points out that we cannot infer its ultimate metaphysical simplicity from this fact. He backs up the point nicely in his 'Refutation of Mendelssohn' (1787: B413-15 and n.): even if a mind or soul is an ultimately metaphysically single

substance, he argues, we can imagine its powers being half what they actually are (its 'degree of reality' being half what it actually is), and this is sufficient to show the sense in which a mind or thinking subject that is strongly unified or simple *may nonetheless be composite or have parts*.

- 43. Perhaps the best move at this point is Coleman's (see e.g. this volume). Note, though, that the kind of phenomenally qualitied nonexperiential reality he posits is very far from any standard conception of nonexperiential being—and that the problem posed by No Radical Emergence remains as acute as ever.
- 44. See also *De Anima* 425b12–25. It is I believe a mistake to read the explicitly propositional formulation 'aware *that* we are aware' in such a way that the claim isn't also true of nonhuman animals.
- 45. See e.g. Zahavi (1999; 2005), Kriegel and Williford (2006). I try to characterize it in Strawson (2013).
- 46. On this last point see James (1890, ch. 6). Panpsychism also solves what some see as a major problem for the theory of evolution—the problem of why experience evolved at all. (Nietzsche expresses this problem vividly at the beginning of §354 of *The Gay Science*, although I don't agree with the solution he proposes later in the paragraph.)

Bibliography

Aristotle. Nicomachean Ethics. Trans. James Thomson. London: Penguin, c 340 BCE/1953.

- Aristotle. De Anima. Trans. Walter Hett. Cambridge: Harvard University Press, c 340 BCE/ 1936.
- Berkeley, George. "Philosophical Commentaries." *Philosophical Works of George Berkeley*. Ed. Michael Ayers. London: Everyman, 1707–1710/1975. 305–412.
- Bohm, David. Causality and Chance in Modern Physics. London: Routledge, 1957/2005.
- Chalmers, David. "The Combination Problem for Panpsychism." *Panpsychism.* Ed. Godehard Brüntrup and Ludwig Jaskolla. Oxford: Oxford University Press, this volume.
- Clarke, Samuel. The Clarke-Collins Correspondence of 1707–18. Ed. William Uzgalis. Peterborough: Broadview Press, 1707–1718/2011.
- Coleman, Sam. "Being Realistic: Why Physicalism May Entail Panexperientialism." *Consciousness and Its Place in Nature*. Ed. Anthony Freeman. Thorverton: Imprint Academic, 2006. 40–52.
- Coleman, Sam. "Panpsychism and Neutral Monism: How to Make up One's Mind." *Panpsychism*. Ed. Godehard Brüntrup and Ludwig Jaskolla. New York: Oxford University Press, this volume.
- Collins, Antony. *The Clarke-Collins Correspondence of 1707–18*. Ed. William Uzgalis. Peterborough: Broadview Press, 1707–1718/2011.
- Descartes, René. Meditations and Objections and Replies. Rpt. The Philosophical Writings of Descartes, Vol. 2. Cambridge: Cambridge University Press, 1641/1985.
- Eddington, Arthur. "Review of 'The Analysis of Matter." *The Journal of Philosophical Studies* 3.9 (1928): 93–95.
- Faraday, Michael. "A Speculation Touching Electric Conduction and the Nature of Matter." *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 24 (1844): 136–44.
- Foster, John. The Case for Idealism. London: Routledge, 1982.
- Frauenstädt, Julius. *Studien und Kritiken zur Theologie und Philosophie*. Berlin: Verlag der Voss'schen Buchhandlung, 1840.
- Goff, Philip. "Experiences Don't Sum." *Consciousness and Its Place in Nature*. Ed. Anthony Freeman. Thorverton: Imprint Academic, 2006. 53–61.

- Goff, Philip. "The phenomenal bonding solution to the combination problem." *Panpsychism*. Ed. Godehard Brüntrup and Ludwig Jaskolla. New York: Oxford University Press, this volume.
- Griffin, Donald. Unsnarling the World-Knot: Consciousness, Freedom, and the Mind-Body Problem. Berkeley: University of California Press, 1998.
- Hawking, Stephen. A Brief History of Time. New York: Bantam Books, 1988.
- Heisenberg, Werner. Physics and Philosophy. New York: Harper and Brothers, 1958.
- Horgan, Terence and Matjaz Potrč. Austere Realism: Contextual Semantics Meets Minimal Ontology (Cambridge, MA: MIT Press, 2008.
- Jackson, Frank. (2003) "From H₂O to Water: The Relevance to A Priori Passage." *Real Metaphysics*. Ed. Halvard Lillehammer and Gonzalo Rodriguez-Pereyra. London: Routledge, 2003. 84–97.
- James, William. The Principles of Psychology. New York: Dover, 1890/1950.
- Kant, Immanuel. Critique of Pure Reason. London: Macmillan, 1781–87/1933.
- Kriegel, Uriah and Kenneth Williford. "Introduction." Self-Representational Approaches to Consciousness. Ed. Uriah Kriegel and Kenneth Williford. Cambridge, MA: MIT Press, 2006. 1–8.
- Ladyman, James et al. Every Thing Must Go: Metaphysics Naturalized. Oxford: Clarendon Press, 2007.
- Leibniz, Gottfried. *New Essays on Human Understanding*. Trans. Jonathan Bennett and Peter Remnant. Cambridge: Cambridge University Press. *c* 1704/1996.
- Leibniz, Gottfried. *Monadology. Monadology and Other Philosophical Essays.* Transl. Paul Schrecker and Anne Martin Schrecker. Indianapolis: Bobbs-Merrill, 1714/1965.
- Levine, Joseph. "Materialism and Qualia: The Explanatory Gap." *Pacific Philosophical Quarterly* 64.3 (1983): 354–61.
- Lewis, David. Philosophical Papers. Volume II. Oxford: Oxford University Press, 1986.
- Lockwood, Michael. "The Grain Problem." *Objections to Physicalism*. Ed. Howard Robinson. New York: Oxford University Press, 1993. 271–91.
- Lockwood, Michael. "Consciousness and the Quantum World: Putting Qualia on the Map" Consciousness: New Philosophical Perspectives. Ed. Quentin Smith and Aleksandar Jokic. Oxford: Oxford University Press, 2003. 447–67.
- Lucretius. The Nature of Things. Transl. A. Stallings. London: Penguin, c50 bce/2007.
- Newman, Max. "Mr. Russell's 'Causal Theory of Perception'" Mind 37 (1928): 137-48.
- Nietzsche, Friedrich. The Gay Science. Cambridge: Cambridge University Press, 1882/2001.
- Perry, John. Thought and Character of William James. Boston: Little, Brown, 1935.
- Plato. *The Sophist*. Trans. H. N. Fowler. London: Heinemann, *c* 360 BCE/1921.
- Poincaré, Henri. Science and Hypothesis. London and Newcastle-upon-Tyne: Walter Scott, 1903/1905.
- Quine, Willard van Orman. "Two Dogmas of Empiricism." *From a Logical Point of View*. 2nd ed. New York: Harper and Row, 1951/1961.
- Robinson, Howard. "Idealism." *The Oxford Handbook to the Philosophy of Mind*. Ed. Brian McLaughlin, Ansgar Beckermann, and Sven Walter. Oxford: Oxford University Press, 2009.
- Rosenberg, Gregg. "On the Intrinsic Nature of the Physical." *Towards a Science of Consciousness III.* Ed. Stuart Hameroff, Alfred Kaszniak, and David Chalmers. Cambridge: MIT Press, 1999.
- Rosenberg, Gregg. A Place for Consciousness: Probing the Deep Structure of the Natural World. New York: Oxford University Press, 2004.
- Russell, Bertrand. The Analysis of Mind. London: George Allen and Unwin, 1921/1978.
- Russell, Bertrand. The Analysis of Matter. London: Routledge, 1927/1992.
- Russell, Bertrand. "Mind and Matter." *Portraits from Memory*. Nottingham: Spokesman, 1956/1995. 145–65.

- Schaffer, Jonathan. "From Nihilism to Monism." Australasian Journal of Philosophy 85.2 (2007): 175–91.
- Schaffer, Jonathan. "Monism: The Priority of the Whole." *Philosophical Review* 119 (2010): 31–76.
- Schlick, Moritz. General Theory of Knowledge. New York: Springer-Verlag, 1918–25/1974.
- Seager, William. "Panpsychism, Aggregation and Combinatorial Infusion." *Mind & Matter* 8.2 (2010): 167–84.
- Seager, William. *Natural Fabrications*. Dordrecht: Springer, 2012.
- de Spinoza, Baruch. Short Treatise on God, Man and his Well-being. Repr. The Collected Works of Spinoza. Ed. Edwin Curley. Princeton: Princeton University Press, 1662/1988.
- Sprigge, Timothy. "Is Consciousness Mysterious?" *The Importance of Subjectivity*. Ed. Leemon McHenry. Oxford: Oxford University Press, 1999/2010.
- Stoljar, Daniel. "Comments on Galen Strawson" *Consciousness and Its Place in Nature.* Ed. Anthony Freeman. Thorverton: Imprint Academic, 2006a. 170–76.
- Stoljar, Daniel. Ignorance and Imagination. Oxford: Oxford University Press, 2006b.
- Strawson, Galen. "Real Materialism." *Real Materialism and Other Essays*. Oxford: Oxford University Press, 2003a/2008. 19–52.
- Strawson, Galen. "What Is the Relation Between an Experience, the Subject of an Experience, and the Content of the Experience?" *Real Materialism and Other Essays*. Oxford: Oxford University Press, 2003b/2008. 151–88.
- Strawson, Galen. "Realistic Monism: Why Physicalism Entails Panpsychism." Real Materialism and Other Essays. Ed. Galen Strawson. Oxford: Oxford University Press, 2006/2008.53–74.
- Strawson, Galen. "The Identity of the Categorical and the Dispositonal." *Analysis* 68.4 (2008): 271–82.
- Strawson, Galen. "Self-Intimation." Phenomenology and the Cognitive Sciences 14 (2013/ 2015): 44–73.
- Strawson, Galen. "Consciousness in the 20th Century." *Consciousness*. Ed. Alison Simmons. Oxford: Oxford University Press, *forthcoming a*.
- Strawson, Galen. "Structure and Quality." Forthcoming b.
- Turausky, Keith. "Picturing Panpsychism: New Approaches to the Combination Problem(s)." Unpublished.
- Weinberg, Steven. "Before the Big Bang." The New York Review of Books 12 Jun. 1997.
- Whitehead, Alfred. *Process and Reality: An Essay in Cosmology*. Ed. David Griffin and Donald Sherburne. New York: Free Press, 1929/1978.
- Wilson, Rob. *Boundaries of the Mind: The Individual in the Fragile Sciences*. Cambridge: Cambridge University Press, 2004.
- Zahavi, Dan. Self-Awareness and Alterity: A Phenomenological Investigation. Evanston, IL: Northwestern University Press, 1999.
- Zahavi, Dan. Subjectivity and Selfhood: Investigating the First-Person Perspective. Cambridge: MIT Press, 2005.

4

Panpsychism and Priority Cosmopsychism

YUJIN NAGASAWA AND KHAI WAGER

4.1 Introduction

A contemporary form of panpsychism says that phenomenality is prevalent because all physical ultimates instantiate phenomenal or protophenomenal properties. According to priority cosmopsychism, an alternative to panpsychism that we propose in this chapter, phenomenality is prevalent because the whole cosmos instantiates phenomenal or protophenomenal properties. It says, moreover, that the consciousness of the cosmos is ontologically prior to the consciousness of ordinary individuals like us. Since priority cosmopsychism is a highly speculative view our aim in this chapter remains modest and limited. Instead of providing a full defense of priority cosmopsychism. This, however, by no means entails that we develop the view in logical space merely for its own sake. We offer instead a blueprint for a new alternative to panpsychism and explain how such a view avoids some of the most persistent problems for panpsychism while maintaining several of its strengths.¹

This chapter has the following structure. In section 4.2, we discuss panpsychism and priority monism, which are relevant to priority cosmopsychism. In section 4.3, we introduce priority cosmopsychism. In section 4.4, we show that priority cosmopsychism overcomes the main difficulties for panpsychism, including the problem of infinite decomposition and the combination problem. In section 4.5, we defend priority cosmopsychism against possible objections. Section 4.6 concludes.

4.2 Panpsychism and Priority Monism

Priority cosmopsychism is structurally parallel to both panpsychism and priority monism. We therefore address each of these views before formulating priority cosmopsychism.

4.2.1 Panpsychism

Since the present volume is devoted to panpsychism, we will not provide a comprehensive overview of panpsychism here. Nevertheless, some essential preliminaries are in order. The most straightforward version of panpsychism is formulated in terms of ordinary mental states. It says that everything has mental states in the same sense as we do—for example, rocks have thoughts to the same extent that we do. This is highly implausible. Contemporary panpsychism is, on the other hand, typically formulated in terms of phenomenal or protophenomenal properties instead of all types of mental states. There are many contemporary formulations, but in this chapter we focus on Philip Goff's formulation as follows (see Goff 2009, 294):

Panpsychism: All physical ultimates instantiate phenomenal properties.

As Goff notes, this view is closely related to the following view:

Micropsychism: Some physical ultimates instantiate phenomenal properties.

Panpsychism is an extreme form of micropsychism because it says that all, not merely some, physical ultimates instantiate phenomenal properties. That is why the view is called *pan*psychism.

Some formulate panpsychism in terms of protophenomenal properties instead of phenomenal properties. They say that some physical ultimates instantiate protophenomenal, rather than phenomenal, properties. David Chalmers addresses the distinction between the phenomenal and protophenomenal versions of panpsychism:²

There are two ways this might go. Perhaps we might take [phenomenal] experience itself as a fundamental feature of the world, alongside space-time, spin, charge and the like. That is, certain phenomenal properties will have to be taken as *basic* properties. Alternatively, perhaps there is some *other* class of novel fundamental properties from which phenomenal properties are derived.... [T]hese cannot be physical properties, but perhaps they are nonphysical properties of a new variety, on which phenomenal properties are logically supervenient. Such properties would be related to experience in the same way that basic physical properties are related to nonbasic properties such as temperature. We could call these properties *protophenomenal* properties, as they are not themselves phenomenal but together they can yield the phenomenal. (Chalmers 1996, 126–27)

The main reason for holding panpsychism is that it avoids the problem of strong emergence. This problem arises from the unexpectedness of phenomenal properties: phenomenal properties are instantiated by physical entities such as aggregates of neurons, but this is unexpected and surprising because neurons seem to be fundamentally nonexperiential. It seems impossible to explain how something experiential can be instantiated by something fundamentally non-experiential. According to Galen Strawson, the instantiation of experiential phenomena by wholly nonexperiential phenomena is as extraordinary as the instantiation of spatial phenomena by nonspatial phenomena. He contends that such emergences are impossible because the following is true: For any feature Y of anything that is correctly considered to be emergent from X, there must be something about X and X alone in virtue of which Y emerges, and which is sufficient for Y. Strong emergence violates such a law and, hence, it is, "by definition, a miracle every time it occurs" (Strawson 2008, 64–65).

Panpsychism avoids the problem of strong emergence by stipulating that physical ultimates are themselves phenomenal or protophenomenal. That is, according to panpsychism, it is not surprising that phenomenal properties are instantiated by aggregates of neurons because physical ultimates, which constitute neurons and other physical entities, are already phenomenal or protophenomenal.

4.2.2 Priority Monism

Priority monism says that exactly one *basic* concrete object, that is, the cosmos, exists (see Schaffer 2008). Priority monism should be distinguished from existence monism, according to which exactly one concrete object, that is, the cosmos, exists.³ Unlike existence monism, priority monism is compatible with the existence of multiple concrete objects because it says only that there is exactly one *basic* concrete object. According to priority monism, the cosmos is more basic than other concrete objects in the sense that it is onto-logically prior to, or ontologically more fundamental than, those other objects.

In other words, all concrete objects, except the cosmos itself, are derivative of the cosmos.

Priority monism appears counterintuitive initially because in most instances we think that a whole is not ontologically prior to its parts. We think, for example, that the grains of sand constituting a heap are prior to the heap or that tiles in a mosaic are prior to the mosaic. Jonathan Schaffer points out, however, that there are many other examples in which we think that a whole is, in fact, prior to its parts. For instance, we think that a circle is prior to semicircles of the circle or that a body is prior to organs of the body (see Schaffer 2008). This is because, according to Schaffer, our common sense distinguishes between mere heaps and genuine unities. A heap of grains of sand and a mosaic are mere heaps but a circle, a body, and the cosmos are, according to Schaffer, genuine unities.

Schaffer notes that priority monism is concerned with concrete objects and excludes everything else. He writes:

I assume that there is a maximal actual concrete object—*the cosmos* of which all actual concrete objects are parts. I should emphasize that I am only concerned with actual concrete objects. *Possibilia, abstracta,* and actual concreta in categories other than *object* are not my concern (deities and spirits, if such there be, are not my concern either). When I speak of the world—and defend the monistic thesis that the whole is prior to its parts—I am speaking of the material cosmos and its planets, pebbles, particles, and other proper parts. (Schaffer 2010, 33)

Phenomenal properties are not within the scope of priority monism as they are not concrete objects.

4.3 Priority Cosmopsychism

We are now ready to formulate priority cosmopsychism. Again, priority cosmopsychism is structurally parallel to both panpsychism and priority monism.

Consider, first, the parallel structure between priority monism and priority cosmopsychism. Priority monism says that exactly one basic concrete object, the cosmos, exists. In parallel to this, priority cosmopsychism says that exactly one basic consciousness, the cosmic consciousness, exists. Recall that priority monism is concerned only with concrete objects. Priority cosmopsychism is, on the other hand, concerned only with phenomenal and protophenomenal properties, which fall outside the scope of priority monism. Priority cosmopsychism should be distinguished from existence cosmopsychism, according to which exactly one consciousness, the cosmic consciousness, exists. Unlike existence cosmopsychism, priority cosmopsychism is compatible with the existence of multiple individual consciousnesses because it says only that there is exactly one *basic* consciousness. The cosmic consciousness is more basic than other consciousnesses in the sense that it is ontologically prior to or ontologically more fundamental than other consciousnesses. All consciousnesses except the cosmic consciousness itself are derivative of the cosmic consciousness, in a manner similar to that in which all concrete objects except the cosmos itself are, according to priority monism, derivative of the cosmos.

Consider now the parallel structure between panpsychism and priority cosmopsychism. Panpsychism says, again, that all physical ultimates—that is, physical entities on the bottom level of reality—instantiate phenomenal properties. In parallel to this, priority cosmopsychism says that the cosmos, which is on the top level of reality, instantiates phenomenal properties. Panpsychism claims that phenomenal properties that physical ultimates instantiate are more fundamental than phenomenal properties of ordinary individuals. In fact, according to panpsychism, phenomenal properties of physical ultimates are the most fundamental form of phenomenality. In parallel to this claim, priority cosmopsychism says that phenomenal properties which the cosmos instantiates are more fundamental than phenomenal properties of ordinary individuals. In fact, according to priority cosmopsychism, the cosmic consciousness is the most fundamental form of phenomenality.

It is interesting to note that the combination of priority monism and (priority) cosmopsychism entails a unique version of panpsychism. Recall the formulation of panpsychism we adopt in this chapter: all physical ultimates instantiate phenomenal properties. Priority monism says that the phrase 'physical ultimates' in the formulation refers to a single entity, the cosmos, and (priority) cosmopsychism says that the cosmos instantiates phenomenal properties. This means that the combination of priority monism and (priority) cosmopsychism entails that the physical ultimate instantiates phenomenal properties, which is exactly what panpsychism says. In this chapter, however, in order to avoid confusion, by the term 'physical ultimates' we mean fundamental physical entities on the bottom level of reality, rather than the cosmos. Also, we remain neutral about the compatibility of priority monism with (priority) cosmopsychism because priority cosmopsychism does not rely on priority monism (and vice versa). We also remain neutral about the nature of the cosmic consciousness. Some pantheists or panentheists might think that the cosmic consciousness is the consciousness of a higher being, such as God, which shares phenomenal experiences of individual conscious beings. Some others might think that the cosmic consciousness is not in itself phenomenal but only protophenomenal. However, these issues are not crucial to our discussion.

4.4 Priority Cosmopsychism Versus Panpsychism

Why should we consider priority cosmopsychism as a serious alternative to panpsychism? First, like panpsychism, priority cosmopsychism is not vulnerable to the problem of strong emergence. This is because priority cosmopsychism rejects the claim that something experiential can be instantiated by something fundamentally nonexperiential. Second, more importantly, priority cosmopsychism avoids some of the most persistent problems for panpsychism albeit that priority cosmopsychism is structurally parallel to panpsychism. In this section, we consider two such problems, the problem of infinite decomposition and the combination problem.

4.4.1 The Problem of Infinite Decomposition

Again, panpsychism holds that all physical ultimates instantiate phenomenal properties. This means that panpsychism presupposes fundamentalism. Fundamentalism identifies entities on the bottom, fundamental level as ultimate reality.⁴ Consider, for instance, physicalism as a version of fundamentalism. According to one form of physicalism, the ultimate level of reality is physical because, roughly speaking, microphysical theory describes the properties and behaviors of fundamental subatomic particles, on which everything else in the actual world supervenes. This means that entities on the fundamental level are entirely physical and, hence, everything in the actual world is ultimately physical. Panpsychism, at least the version that we have been considering here, adds to this form of physicalism that the fundamental subatomic particles, that is, physical ultimates, instantiate phenomenal properties.

Schaffer (2003) and Montero (2006) consider the argument that physicalism is false because fundamentalism is false. According to this argument, since the cosmos is stratified infinitely into levels, physicalism cannot be true. They are right in thinking that, insofar as physicalism is formulated as a version of fundamentalism, the falsity of fundamentalism entails the falsity of physicalism. However, the falsity of fundamentalism also entails the falsity of panpsychism because, again, panpsychism presupposes fundamentalism.

Schaffer tries to show that it is at least possible that the cosmos is stratified infinitely into levels by appealing to the conceivability and logical consistency of infinite decomposition (see Schaffer 2003, 501). First, he says, infinite decomposition is metaphysically possible because it is conceivable that everything has parts. It is conceivable that everything is extended and everything that is extended is decomposed into further entities. If conceivability entails

possibility, then it is possible that everything has parts. Second, he says, infinite decomposition is metaphysically possible because it is logically consistent. There are consistent models of mereology that allow infinite decomposition. Given that there are such consistent models, there is no a priori ground for rejecting the possibility of infinite decomposition as a metaphysical possibility. Schaffer contends, moreover, that infinite decomposition might be not only possible but also actual because it is taken seriously by scientists. For example, the quantum physicist David Bohm (1957) says that his formulation of physics is "consistent with an infinity of levels." To take another example, the physicist Hans Dehmelt (1989) postulates an infinite regression of subelectron structure. So it appears that while it remains inconclusive whether the lack of physical ultimates is actually true, it should be taken seriously.

Again, if fundamentalism is false and there are no physical ultimates, then panpsychism is false. In such a case, contrary to what panpsychism says, there are no physical ultimates to instantiate phenomenal properties. One might suggest at this point that if there are no physical ultimates, then panpsychism can be defined as a thesis that certain microphysical entities, but not physical ultimates, instantiate phenomenal properties. However, such a view is arbitrary. It is unclear why certain microphysical entities on a certain level of reality instantiate phenomenal properties while others on lower levels do not. The possibility of infinite decomposition therefore threatens panpsychism.

Priority cosmopsychism, however, is not vulnerable to the problem of infinite decomposition. This is because priority cosmopsychism does not rely on fundamentalism. More specifically, it attributes basic consciousness to the cosmos, which is on the top level of reality, rather than physical ultimates, which, if they exist, are on the bottom level. Whether there is a bottom level, therefore, is irrelevant to the cogency of priority cosmopsychism. As long as the cosmos exists, priority cosmopsychism is intact, and indeed the cosmos does exist. These observations give us a reason to prefer priority cosmopsychism to panpsychism.

We have considered the possibility of infinite decomposition of concrete objects, but we might extend this idea to phenomenal properties as well. Chalmers, for example, seems to think that phenomenal properties are properly arranged sums of protophenomenal properties when he says that phenomenal properties logically supervene on protophenomenal properties (see Chalmers 1996, 126). If that is true, it might be the case that phenomenal properties are infinitely decomposable into more and more primitive forms of protophenomenal properties and that the chain of decomposition or supervenience continues infinitely. Such a possibility would also undermine panpsychism because the whole point of panpsychism is to introduce phenomenal or protophenomenal properties as fundamental building blocks of phenomenal reality on the bottom level so that the existence of consciousness does not entail strong emergence. If phenomenal properties are infinitely decomposable they cannot be fundamental building blocks.

Priority cosmopsychism is not threatened by the possibility of infinite decomposition of phenomenal properties either, because, again, priority cosmopsychism regards the cosmic consciousness as ontologically prior to 'smaller' forms of consciousness, so whether there are 'smallest' forms of phenomenal or protophenomenal properties is irrelevant to the cogency of priority cosmopsychism.⁵

4.4.2 The Combination Problem

The combination problem arises from the apparent discrepancy between a highly complex, structured aggregate of atoms and brain cells, on the one hand; and a smooth, uniform phenomenal experience such as a visual experience, on the other. The problem can be formulated as an objection to panpsychism as follows: Ordinary phenomenal experiences present themselves as smooth, continuous, and unified. They do have distinct aspects, but they have an underlying homogeneity. According to panpsychism, however, all physical ultimates instantiate phenomenal or protophenomenal properties and our ordinary phenomenal experiences result from combinations of these properties. It is hard to see, however, how phenomenal or protophenomenal properties of microphysical entities could add up to the homogeneous character of phenomenal experiences that we have.

The combination problem is arguably the most difficult problem for panpsychism. Chalmers, for example, writes, "It is certainly the hardest problem for any sort of Russellian view" (which includes a version of panpsychism we consider here; Chalmers 1996, 307). William Seager also regards it as "the most difficult problem facing any panpsychist theory of consciousness" (Seager 1995, 280). Priority cosmopsychism, however, does not face the combination problem because, unlike panpsychism, it denies that phenomenal experiences are constituted by phenomenal properties of physical ultimates.⁶ Again, priority cosmopsychism attributes basic consciousness to the cosmos and regards individual consciousnesses as derivatives of it. That is, contrary to what panpsychism says, priority cosmopsychism regards phenomenal experiences as derivatives of something 'larger' (i.e., the cosmic consciousness) rather than as the aggregate of something 'smaller' (i.e., phenomenal or protophenomenal properties of physical ultimates). In other words, panpsychism faces the combination problem because it is a bottom-up view—it starts with phenomenal properties or protophenomenal properties of physical ultimates and tries to build ordinary phenomenal properties from them. Priority cosmopsychism,

on the other hand, is a top-down view—it starts with the cosmic consciousness and tries to derive ordinary phenomenal properties from it. Here is an analogy to illustrate this point. Suppose, *per impossibile*, there is an absolutely perfectly smooth painting, which is analogous to a smooth, homogeneous phenomenal experience. Such a painting cannot be an aggregate of small dots, which are analogous to phenomenal or protophenomenal properties of physical ultimates, but it can be a segment of a larger painting that is equally smooth and homogeneous, which is analogous to the cosmic consciousness.

One might point out here that while priority cosmopsychism avoids the combination problem it does seem to face a problem of the same structure on a larger scale. The combination problem asks how medium-size consciousnesses can be built from minute phenomenal or protophenomenal properties of physical ultimates. Similarly, the problem in question asks how the cosmic consciousness can be built from medium-size individual consciousnesses.

Fortunately, this is not a serious problem because it is based on a misinterpretation of priority cosmopsychism. Priority cosmopsychism says that medium-size individual consciousnesses are derivatives of the cosmic consciousness but that does not entail that medium-size individual consciousnesses constitute the cosmic consciousness as ontologically prior building blocks of the cosmic consciousness. On the contrary, according to priority cosmopsychism, the cosmic consciousness is ontologically prior to mediumsize individual consciousnesses.

One might claim, however, that priority cosmopsychism still fails to provide an answer to the following crucial question: How could medium-size individual consciousnesses be derived from the cosmic consciousness? Let us call this problem the 'derivation problem.' It is not easy to provide an answer to the derivation problem because we do not know the exact nature of the cosmic consciousness. Yet we can speculate how we might be able to respond to the problem.

It is reasonable to assume that the cosmic consciousness is somewhat comparable to the consciousness of an ordinary individual because, after all, it is a form of consciousness. If we can then show that the consciousness of an ordinary individual can be divided into smaller, less fundamental segments, then we have reason to think that the cosmic consciousness can also be divided into smaller, less fundamental segments. And it seems indeed possible to divide the consciousness of an ordinary individual into smaller segments.

Consider, for example, a visual experience. A visual experience can be considered to be a unity which may be segmented into distinguishable color experiences (e.g., experiences corresponding to red and green hues) or experiences of separable regions in space (e.g., experiences corresponding to the right-hand side and the left-hand side of the visual field). Yet the whole visual experience is considered to be a unity that is more fundamental than the segments. Perhaps the cosmic consciousness unifies individual consciousnesses in a similar way. The cosmic consciousness is more fundamental than individual consciousness, so it is not the case that individual consciousnesses are fundamental building blocks of the cosmic consciousness. On the contrary, smooth, continuous and unified individual consciousnesses are derived from the smooth, continuous and unified cosmic consciousness.

It may be useful to recall, here, that priority cosmopsychism shares a parallel structure with priority monism. Priority monism states that the concrete cosmos, as an integrated whole, is the only basic concrete object and other ordinary concrete objects are derived from it. Priority cosmopsychism states that the cosmic consciousness, as an integrated whole, is the only basic form of consciousness and ordinary consciousnesses are derived from it. As a result of this parallel structure, just as priority cosmopsychism has to address the derivation problem, so too priority monism has to address its own equivalent of the derivation problem. In the case of priority monism, the derivation problem can be stated as the problem of how the many concrete parts of the cosmos are derived from the basic concrete whole.

Schaffer (2010, 57) offers a number of possible solutions to the derivation problem for priority monism and the same responses can be adapted to answer the derivation problem for priority cosmopsychism. As such, priority cosmopsychism can offer accounts of how the derivation problem might be resolved.

Recall that for priority monism the derivation problem is the problem of accounting for the derivative parts in terms of the basic cosmos. Schaffer addresses the problem in terms of heterogeneity. It is typically an uncontroversial premise that the basic feature(s) of the cosmos must be homogenous. According to priority monism, the cosmos itself is the only basic feature, yet it claims that the cosmos is also heterogeneous because it contains derivative parts. Schaffer offers three different options for explaining the heterogeneity of the cosmos whilst still allowing that it is, as an integrated whole, basic. He also notes that *any* view positing basic features needs to account for their being heterogeneous as opposed to homogenous (see Schaffer 2010). The three accounts of the heterogeneity of the cosmos are given with respect to firstly, distributional properties, secondly, regionalized properties, and finally, regionalized instantiation.

On the first account the cosmos, as an integrated whole, is heterogeneous due to instantiating distributional properties,

For the monist, the general fact that the world is heterogeneous is due to the world's instantiating the determinable property of being heterogeneous. The specific way that the world is heterogeneous is due to the world's instantiating the determinate property of tracing suchand-such a curve through physical configuration space. Thus the one whole can be parturient. (Schaffer 2010, 260)

On the second account, the cosmos is heterogeneous due to regionalized properties. The cosmos has the monadic property of being the cosmos, yet it bears a relation of, say, spikiness to one region and flatness to another. The third account also makes use of regionalization but instead appeals to regionalized instantiation, where the cosmos is heterogeneous due to it, say, instantiatinghere spiky and instantiating-there flat.

There are differing views regarding the three accounts, but the important thing is that they are consistent ways to make the move from, in concrete terms, a cosmos that is a basic integrated whole to a derivative heterogeneity. As a result of priority cosmopsychism sharing a parallel structure with priority monism, we might adopt these strategies in response to the derivation problem for priority cosmopsychism. A version of all three accounts could be given to explain the heterogeneity of the cosmic consciousness.

In parallel to the first response, priority cosmopsychists might say that the cosmic consciousness is heterogeneous due to it instantiating the determinable property of being heterogeneous. According to this response, the cosmic consciousness would instantiate the distributive property of following a particular path through phenomenal configuration space (no doubt an extremely complex path through a configuration space of many dimensions). In parallel to the second response, priority cosmopsychists might say that the cosmic consciousness is heterogeneous due to regionalized properties, where the cosmic consciousness is a monadic property which bears a relation of redness to one region and blueness to another region. The monadic property of being the cosmic consciousness would demonstrate many relations among regionalized phenomenal properties. Finally, in parallel to the third response, priority cosmopsychists might say that the cosmic consciousness is heterogeneous due to regionalized instantiation of phenomenal properties, the cosmic consciousness instantiates-here red and instantiates-there blue. A thorough exploration of such possibilities is not within bounds of the present chapter but will make for interesting future work.

Let us recap. Panpsychism faces the problem of infinite decomposition because it relies on fundamentalism. Priority cosmopsychism, on the other hand, does not face that problem as it is free from fundamentalism. Panpsychism also faces the combination problem, which is recognized as the strongest objection to the view. Priority cosmopsychism, on the other hand, offers a satisfactory answer to this problem. Instead of the combination problem, however, priority cosmopsychism faces the derivation problem. Yet, as we have seen, there are prima facie reasons to think that it can be resolved. Therefore, priority cosmopsychism seems more attractive than panpsychism.

4.5 Objections to Priority Cosmopsychism

We have seen that priority cosmopsychism overcomes some of the most persistent problems associated with panpsychism. One might argue, however, that priority cosmopsychism still seems more implausible than panpsychism. In this section, we review some objections to priority cosmopsychism.

4.5.1 Inexplicability of the Cosmic Consciousness

One might reject priority cosmopsychism by saying that it is silent about exactly what the cosmic consciousness is. The attribution of phenomenality to the cosmos is essential for priority cosmopsychism, so without explaining what the cosmic consciousness is, one might say, priority cosmopsychism is incomplete.

Priority cosmopsychism is not completely silent about the nature of the cosmic consciousness. It says, for example, that the cosmic consciousness is ontologically the most fundamental form of consciousness of which the consciousnesses of ordinary individuals are derivative. We can also speculate about further possibilities. For example, we might think that since the cosmos on the whole is not complex enough in a relevant sense to instantiate phenomenality to the fullest extent there is no such thing as the phenomenal self for the cosmic consciousness. Perhaps the cosmic consciousness is an organic unity of phenomenal and protophenomenal forms of conscious experiences. Recall again however that our purpose here is not to offer a full defense of priority cosmopsychism but only to show that priority cosmopsychism is more attractive than panpsychism insofar as it avoids some of the most persistent problems for panpsychism. If panpsychism does not say much about the nature of the consciousness of physical ultimates, priority cosmopsychism is not committed to saying much about the nature of the cosmic consciousness either. And, in fact, panpsychism says very little about the consciousness of physical ultimates. Chalmers, for example, writes, "Of course it is very hard to imagine what a protophenomenal property [which a physical ultimate instantiates] could be like but we cannot rule out the possibility that they exist" (Chalmers 1996, 127). We can make a parallel claim here: Of course it is very hard to imagine what the cosmic consciousness could be like but we cannot rule out the possibility that it exists. And, again, there are reasons to prefer priority cosmopsychism to panpsychism.

4.5.2 Counterintuitiveness

Priority cosmopsychism attributes consciousness to the cosmos, which seems highly counterintuitive. One might wonder how we could take such a counter-intuitive thesis seriously.

Recall, once again, that we are comparing only the plausibility of priority cosmopsychism with that of panpsychism. So our interest here is to show only that priority cosmopsychism is no more counterintuitive than panpsychism. Panpsychism holds the fundamentalist view that there is a fundamental bottom level of reality and it adds that physical ultimates on the fundamental level instantiate phenomenal properties. Priority cosmopsychism, on the other hand, holds that the cosmos is on the top level of reality and adds that the cosmos instantiates phenomenal properties. Structurally speaking, therefore, they are parallel, and there seems no reason to think that either of them is distinctively more counterintuitive than the other.

One might claim, however, that the attribution of phenomenality to the cosmos is particularly absurd. The brain can instantiate phenomenal properties because it has the right structural complexity. Yet, one might continue, the cosmos is not comparable to the brain in terms of structural complexity.

While this might be a good argument to show that priority cosmopsychism is counterintuitive it is not a good argument to show that priority cosmopsychism is *more* counterintuitive than panpsychism. This is because panpsychism faces an objection of the exact same form: Physical ultimates do not have the structural complexity of the brain, so it is counterintuitive to think that they can instantiate phenomenal properties. (If structural complexity is really crucial it might be more implausible to say that physical ultimates have consciousness than that the cosmos does because they are structurally much less complex than the cosmos.)

Notice that panpsychism itself is often rejected on the ground that it is highly counterintuitive. John Searle, for example, calls panpsychism an "absurd view" and characterizes Chalmers's defense of panpsychism as follows: "when faced with a *reductio ad absurdum* argument he just accepts the absurdity" (Searle 1997, 156). It would be ironic if panpsychists were to dismiss priority cosmopsychism because of its counterintuitiveness when they emphasize that panpsychism should not be dismissed on the basis of its counterintuitiveness.

We can apply the same reasoning to many other objections to priority cosmopsychism. For example, one might say that priority cosmopsychism is absurd because there is no sign that the cosmos is conscious (the 'no sign' problem for priority cosmopsychism) or because there is no definitive empirical test to prove that the cosmos is conscious (the 'no test' problem for priority cosmopsychism). In response to the 'no sign' problem, one might say that there is no sign because the cosmos is not structured in such a way that it behaves in accordance with the phenomenal or protophenomenal properties it has, unlike the way in which human bodies behave in accordance with the phenomenal or protophenomenal properties humans have. In response to the 'no test' problem, one might point out that, to the extent that there is no definitive empirical test to prove that the cosmos has consciousness, there is similarly no definitive empirical test to prove that higher animals have consciousness. That is why the problem of animal minds (and other minds) is intractable. However, it is unnecessary to offer such philosophically substantial responses because these problems apply as much to panpsychism as to priority cosmopsychism. There is no sign that physical ultimates are conscious (the 'no sign' problem for panpsychism) and there is no definitive empirical test to prove that physical ultimates are conscious (the 'no test' problem for panpsychism). Again, we are comparing only priority cosmopsychism and panpsychism. It is, therefore, sufficient to say that while these problems might be genuine challenges for priority cosmopsychism they apply equally to panpsychism. Hence, these problems do not make priority cosmopsychism any more implausible than panpsychism.

4.5.3 Estrangement from Current Science

It might be contended that priority cosmopsychism is not to be preferred since it is less compatible with features of current science than contemporary panpsychism is. It might be argued, for example, that priority cosmopsychism is an especially estranged view since it is not concerned with the same physical ultimates that are the focus of current physics. One might claim panpsychism is preferable on the grounds that it is concerned with the same physical ultimates described by current physics, since it states that fundamental phenomenal, or protophenomenal, properties are associated in some sense with such ultimates.

One particular objection of this kind might be that priority cosmopsychism is unable to adhere to the causal closure of the microphysical. This is the principle which says that the causal efficacy of the world is fully accounted for in terms of the causal efficacy of the physical ultimates. One might claim that panpsychism can address the problem of causal closure but priority cosmopsychism cannot. Panpsychism might adhere to the principle by claiming that since all physical ultimates instantiate phenomenal properties any causal efficacy that they may have is already accounted for in current physics.

In response to such objections, we first note that the purpose of this chapter is to defend a blueprint for a new alternative to panpsychism, here we do not defend any specific view based on this blueprint. In this chapter we only address phenomenality and do not endorse a particular relation between phenomenal properties and physical properties. Since it is in such a relation that it will become clear if priority cosmopsychism can adhere to the causal closure of the microphysical, it is after developing a specific view based on the blueprint that one would be fully equipped to respond to this objection. However, it might be interesting to note that one possible development on the blueprint we offer here is a *dual-aspect* version of priority cosmopsychism, according to which the phenomenal and the physical are co-extensive, with the respective properties at the level of the cosmos being basic. On such a view it might be considered more plausible for the priority cosmopsychist to follow the panpsychist in claiming that the principle of causal closure is adhered to on the grounds of the phenomenal already being accounted for in our current physics.

4.6 Conclusion

Panpsychism is an attractive view because, by attributing phenomenality to the fundamental nature of reality, it avoids the problem of strong emergence. However, on the other hand, panpsychism faces the infinite decomposition problem because it presupposes the existence of physical ultimates. It also faces the combination problem because it holds that phenomenal experiences are constituted by phenomenal or protophenomenal properties of physical ultimates. Priority cosmopsychism can be construed as a hypothesis designed to avoid these problems without compromising the promising approach to the problem of strong emergence suggested by panpsychism. Priority cosmopsychism attributes the most fundamental form of consciousness to the cosmos, rather than physical ultimates, and holds that the consciousnesses of ordinary individuals are derivative of it. In this way, priority cosmopsychism avoids not only the problem of strong emergence but also the infinite decomposition problem and the combination problem. Since priority cosmopsychism and panpsychism are structurally parallel, priority cosmopsychism is no more implausible or counterintuitive than panpsychism. Therefore, we can conclude that priority cosmopsychism benefits from a theoretical advantage over panpsychism.

Again, what we have defended in this chapter is a blueprint for a new alternative to panpsychism. This blueprint may be used to develop more specific views, such as monistic, dualistic, or even pantheistic views based on priority cosmopsychism. We have to wait for another occasion to develop and assess such specific views.

Notes

- An earlier version of this essay was presented at the "Minds: Human and Divine" conference in Munich in 2012. This conference was part of the Analytic Theology Project graciously funded by the John Templeton Foundation. We would like to thank all in the audiences. We are particularly grateful to Godehard Brüntrup who organized the event. This essay was written as part of Nagasawa's research project with Andrei Buckareff, "Exploring Alternative Concepts of God," funded by the John Templeton Foundation. We thank the Foundation for its generous support.
- 2. To be precise, in this passage Chalmers is talking about the phenomenal and protophenomenal versions of what he calls type-F monism, which subsumes some versions of panpsychism. So his focus in the passage is more general than ours.
- 3. For a discussion of existence monism see Horgan and Potrč 2000.
- 4. As Barbara Montero (2006, 181) points out, fundamentalism can be formulated in many ways. For example, it can be formulated in terms of decomposition, in which case entities on the fundamental level are undecomposable proper parts (i.e., mereological atoms or simples) that constitute everything else on higher levels. To take another example, it could be formulated in terms of supervenience, in which case entities on the fundamental level are the bases on which all entities on higher levels supervene. It can also be formulated in terms of realization, explanation, reduction, determination, and so on. In this chapter, we focus on decomposition because that seems to be most intuitive. However, most of the claims that we make over the course of this paper apply equally to other formulations.
- 5. Here we use the term 'small' metaphorically. Phenomenal properties are not concrete objects so, of course, they do not occupy physical space.
- 6. Similar points are made by Ludwig Jaskolla and Alexander Buck (2012) and Freya Mathews (2011), but the cosmopsychist views to which they appeal are radically different from ours. Consider, first, Jaskolla and Buck's 'panexperientialist holism.' Panexperiential holism presupposes existence monism, saying "there is exactly one entity-the universe itself" (Jaskolla and Buck 2012, 196). Existence monism is a highly controversial thesis, on which our view, priority cosmopsychism, does not rely. Priority cosmopsychism does not even rely on priority monism, which is more modest than existence monism. Panexperiential holism also stipulates that the universe is "a subject of experience . . . exemplifying experiential content" (Jaskolla and Buck 2012, 196). Priority cosmopsychism does not make such a claim as it is a minimalist view that is parallel to panpsychism. Insofar as panpsychism does not assume that physical ultimates are subjects of experience exemplifying experiential content, priority cosmopsychism does not assume that the cosmos is a subject of experience exemplifying experiential content. Consider, second, Freya Mathews's "cosmological panpsychism." According to this view, "the One" is a subject that "may feel the effects of finite centres of subjectivity in the field of its own larger subjectivity, even though it may not be able actually to experience the way such finite selves feel to themselves" (Mathews 2011, 149). Priority cosmopsychism is not committed to such a claim as, again, it does not assume that the cosmos is a subject of experience. Also, in explaining the nature of the consciousness of the One, Mathews appeals to an idea in psychoanalysis saying, "Amongst the unconscious components of psyche are enduring constellations of psychophysical energy which never surface into ego consciousness yet which nevertheless may be active in the psychic life of a person" (Mathews 2011, 148). Again, priority cosmopsychism does not make such a claim.

Bibliography

Bohm, David. Causality and Chance in Modern Physics. London: Routledge and Kegan Paul, 1957.

Chalmers, David. The Conscious Mind. New York: Oxford University Press, 1996.

- Dehmelt, Hans. "Triton, ... Electron, ... Cosmon... An Infinite Regression?" Proceedings of the National Academy of Sciences 86.22 (1989): 8618–19.
- Goff, Philip. "Why Panpsychism Doesn't Help Us Explain Consciousness." *Dialectica* 63.3 (2009): 289–311.
- Horgan, Terry, and Matjaz Potrc. "Blobjectivism and Indirect Correspondence." Facta Philosophica 2.2 (2000): 249–70.
- Jaskolla, Ludwig, and Alexander Buck. "Does Panexperiential Holism Solve the Combination Problem?" *Journal of Consciousness Studies* 19.9–10 (2012): 190–99.
- Mathews, Freya. "Panpsychism as Paradigm." *The Mental as Fundamental.* Ed. Michael Blamauer. Frankfurt: Ontos, 2011. 141–55.
- Montero, Barbara. "Physicalism in an Infinitely Decomposable World." *Erkenntnis* 64.2 (2006): 177–91.
- Schaffer, Jonathan. "Is There a Fundamental Level?" Noûs 37.3 (2003): 498-17.
- Schaffer, Jonathan. "Monism." *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition). Ed. Edward N. Zalta. Web. 8 Apr. 2016.
- Schaffer, Jonathan. "Monism: The Priority of the Whole." *Philosophical Review* 119.1 (2010): 31-76.
- Seager, William. "Consciousness, Information and Panpsychism." Journal of Consciousness Studies 2.3 (1995): 272–88.
- Searle, John. The Mystery of Consciousness. New York: New York Review of Books, 1997.
- Strawson, Galen. "Realistic Monism: Why Physicalism Entails Panpsychism." *Real Materialism and Other Essays.* Ed. Galen Strawson. Oxford: Oxford University Press, 2008. 53–74.

5

In Search of Mentons

Panpsychism, Physicalism, and the Missing Link

BERIT BROGAARD

5.1 Introduction

One of the most puzzling features of consciousness is the problem David Chalmers has called 'the hard problem,' viz, that of explaining why and how our rich inner life, also known as our 'phenomenal consciousness,' arises from a physical basis (see Chalmers 1995). Since Chalmers initially formulated the hard problem, he has argued extensively against standard versions of reductive and nonreductive physicalism (see Chalmers 1996 and subsequent articles on the conceivability argument). Here I shall take standard versions of reductive physicalism to imply that all truths about consciousness are deducible, in principle, from truths about microphysical properties and the laws of physics. Standard versions of nonreductive physicalism, by contrast, imply that not all truths about consciousness are deducible in this way but that all truths about consciousness metaphysically supervene on, or are metaphysically determined by, truths about microphysical properties and the laws of physics. In this chapter I am not concerned with whether Chalmers's arguments against these positions succeed. I shall simply assume that they do in order to pave the way for an exploration of alternatives.

The most immediately appealing alternatives, in my opinion, are strong emergentism and panpsychism, and those are the views I explore here. Strong emergentism is the view that consciousness is a novel property of brain states that emerges from physical processes when we reach a certain level of organization. Panpsychism is the position that the ultimate constituents of the natural world instantiate microphenomenal or protophenomenal properties. Phenomenal properties are properties that all and only conscious things possess. It is in virtue of possessing phenomenal properties that organisms and mental states have experiences. Protophenomenal properties are properties that, when instantiated in the right kind of way, result in phenomenal properties. The view that the ultimate constituents of the world instantiate protophenomenal properties is also sometimes known as 'panprotopsychism.' I shall, however, use the term 'panpsychism' broadly enough to encompass both positions.

The plan of the chapter is as follows. I start by offering several arguments against strong emergentism. I then sketch a version of constitutive panpsychism that takes primordial consciousness to be composed of elementary particles that instantiate microphenomenal qualities. I argue further that even if the fundamental constituents of consciousness are discoverable by science, this does not vindicate any standard type of reductive physicalism for two reasons. First, the fundamental constituents of consciousness are not likely to be needed to explain purely nonconscious phenomena. Second, acclaimed physical theories rest on the assumption that consciousness is something separate from the remainder of the universe. Whether the resultant view is ultimately a form of reductive physicalism or a version of constitutive panpsychism is probably a verbal issue. In the final part of the chapter I consider one of the main critiques of panpsychism, the 'combination problem,' that is, the problem of explaining how complex conscious states arise from the primitive microexperiences panpsychists ascribe to the ultimate constituents of the universe. I show that the proposed version of panpsychism readily blocks three versions of this problem.

5.2 The Mystery of Strong Emergence

A common, and sometimes dismissive, approach to the hard problem is to say that consciousness emerges from low-level physical properties. When people say that a property emerges at a certain level of complexity, they typically mean that a property that is absent at a lower level appears, or comes into being, at a higher level. The emergent property is novel and unexpected compared with the properties of the emergence base.

There are, however, two very different things that people can mean when they talk about emergent properties. On a weak version, truths about the emergent property are deducible, at least in principle, from the low-level phenomenon. The apparent novelty of a weakly emergent property is an artifact of the limited reasoning skills of mortal human beings. As Chalmers (2006) puts it, weakly emergent properties are interesting, nonobvious features that are interesting and nonobvious to us relative to the perceived simplicity of the underlying principles governing the system. On a strong version, truths about the emergent property are not deducible, even in principle, from the low-level phenomenon. C. D. Broad puts it as follows:

[in the case of strong emergence] the characteristic behaviour of the whole could not, even in theory, be deduced from the most complete knowledge of the behaviour of its components, taken separately or in other combinations, and of their proportions and arrangements in this whole. (Broad 1925, 59)

Weak emergence is uncontroversial. The property of being liquid water, for example, weakly emerges from the low-level properties characterizing hydrogen and oxygen molecules. Weakly emergent properties can be derived, at least in principle, from complete knowledge of microlevel information (see Bedau 1997). Weak emergence is thus perfectly compatible with classical forms of reductive physicalism. So, it is not a candidate view for exploration in this paper.

Unlike weak emergence, strong emergence is both highly controversial and deeply problematic. I will look at three different versions of strong emergence and show how each is problematic. The notion of an intrinsic property plays a crucial role in most formulations of emergence. David Lewis (1986, 59–69) defines intrinsic properties as properties that cannot differ between a pair of duplicate objects (perhaps inhabiting different possible worlds), where duplicate objects are those that have the same perfectly natural properties:

Property P is intrinsic = $_{def}$ Necessarily, for any pair of duplicate objects x and y, x instantiates P just in case y instantiates P

Given this notion of an intrinsic property, we can formulate what I will call 'Microphysical Supervenience' as follows:¹

Microphysical Supervenience (MS): Necessarily, if atoms A_1 through A_n compose a living organism O_1 that exemplifies certain phenomenal properties, then any atoms like A_1 through A_n in all their respective intrinsic properties, related to one another by all the same restricted atom-to-atom relations as A_1 through A_n , compose a living organism O_2 with the same phenomenal properties as O_1 .

Restricted atom-to-atom relations should be understood as spatio-temporal and causal relations. Notice that Microphysical Supervenience does not yield an implausible internalist view of consciousness, as changes in the environment of the organism naturally will lead to changes in the atomic constituents, their intrinsic properties or the restricted atom-to-atom relations.

Given Lewis's definition of 'intrinsic property,' Microphysical Supervenience basically states that organisms that share all natural properties and relations have experiences with the same phenomenal character. The principle is thus consistent with panpsychism, as two organisms consisting of atoms instantiating the same intrinsic properties and standing in the same restricted atom-to-atom relations would instantiate the same microphenomenal properties and hence should instantiate the same macroscopic phenomenal properties. If Microphysical Supervenience is false, then physicalism and panpsychism (that is not combined with emergentism) are false. The phenomenal character of experience is metaphysically primitive or derives from something supernatural.

According to the most common formulations of strong emergentism, emergence implies a negation of Microphysical Supervenience plus some causal principles connecting the emergence base with metaphysically primitive, emergent properties (see Chalmers 1996; O'Connor 2000a; O'Connor 2000b). Though the idea of strong emergence is often cashed out informally in terms of lack of deducibility, this is consistent with a view that rejects Microphysical Supervenience. If the emergence base does not determine the emergent properties, then naturally we cannot deduce truths about an emergent property from truth about the emergence base. The novelty of the emergent properties thus derives at least in part from the lack of determination by the emergence base. This notion of emergence, sometimes called 'ontological emergence' (or perhaps the notion outlined below) appears to have been dominant among the British emergentists (see Broad 1925; see also McLaughlin 1992 and Chalmers 1995), and several contemporaries have considered (and in some cases rejected) this notion (see Van Cleve 1990; McLaughlin 1997; Kim 1999; Van Gulick 2001).

A second, related, way to understand strong emergence maintains that strongly emergent phenomena metaphysically supervene on low-level facts (see O'Connor 1994). On this view, the reason the emergent properties are not deducible from those facts is that these emergent properties are metaphysically primitives in their own right; that is, the emergent properties are novel, fundamental types of properties that are not constituted by more fundamental properties. They are, in this sense, unstructured. This may seem counterintuitive, given that the emergent properties are supposed to metaphysically supervene on the emergence base, which some people would take to suffice for a constitutive relationship. To avoid those connotations, some say that the metaphysical supervenience relation is 'brute' in the sense that its instantiation does not settle any issues regarding constitution, ontological priority or ontological innocence (see McLaughlin and Bennett 2011 for discussion). Another way to think of it is that even if emergent properties are completely distinct from the emergence base, metaphysical supervenience can still obtain as long as there are invariant correlations between the two entities (see Kim 1999; Van Gulick 2001).

On both views, fundamental physical laws need to be supplemented with further fundamental psychophysical 'bridge' or 'supervenience' laws to account for how emergent properties depend on low-level properties. The bridge laws are irreducible to laws characterizing properties at lower levels of complexity. What mandates the fundamental psychophysical laws comes from reflections on how to specify the relationship between the emergence base and the emergent property (see Chalmers 2006). On the first view, the emergent property is thought to be causally related to the emergence base. On the second view, the emergent property stands in a relation of brute supervenience to the emergence base. Either way, the emergent property is not nomologically or metaphysically constituted or determined by the emergence base. That is, the microphysical laws and the microfeatures do not necessitate the macrooutcome. Rather, the macrooutcome is necessitated by the microphysical laws, the microfeatures *and* the psychophysical laws.

Both views can, at least at first glance, easily be combined with the view that consciousness is epiphenomenal, thus avoiding a rejection of the view that the physical domain is causally closed or accepting causal overdetermination (see Chalmers 1996). However, both views could also, and often have been, associated with the view that consciousness can causally affect the physical world in a novel way. Timothy O'Connor (1994), for example, defines strong emergence as follows: Property P is an emergent property of a (mereologically complex) object O iff P strongly supervenes on properties of the parts of O, P is not had by any of the object's parts, P is distinct from any structural property of O, and P has a direct ('downward') determinative influence on the pattern of behavior involving O's parts. The idea that consciousness can causally affect the physical world in a novel way seems to be a natural implication of the causal theory of intentional action, according to which mental states, such as beliefs, desires and intentions, cause our intentional actions.

One problem with these two types of strong emergentism is that, with few exceptions, defenders will insist that the emergent properties supervene on the emergence base. If the supervenience relation is metaphysical, then it is a sheer coincidence—a miracle of sorts—that it obtains in spite of the fact that the emergence base does not necessitate the emergent properties, not even together with the laws of physics. The emergent properties are necessitated by the emergence base, the physical laws *and* the fundamental psychophysical laws. But that makes the metaphysically invariant correlation between the emergence base and the emergence properties utterly mysterious.

If the supervenience relation is nomological (see Chalmers 1996), then duplicate organisms can differ in whether they are conscious in universes with different laws of physics. Yet there is a nomologically invariant correlation between the emergent properties and the emergence base: Duplicate organisms cannot differ in terms of whether they are conscious in universes *with the same laws of physics*. But the emergence base plus the laws of physics do not metaphysically necessitate the emergent properties. They are metaphysically necessitated by the emergence base, the laws of physics *and* the fundamental psychophysical laws. So, unless it's metaphysically impossible for the laws of physics and the microfeatures to obtain without the fundamental psychophysical laws, then the nomologically invariant correlations between emergent property and base is a true (and very mysterious) coincidence.

But it seems to be metaphysically possible for the laws of physics and the microfeatures to obtain without the psychophysical laws. Of course, whether it is metaphysically possible will depend on whether the laws of physics necessitate the psychophysical laws. But strong emergentists normally deny that this is the case (see Chalmers 1996). On this view, the physical laws can obtain without the psychophysical laws. But even in worlds where the physical laws obtain but the psychophysical laws do not, the emergence base and the emergent property both obtain despite the fact that the first (together with the laws of physics) does not necessitate the second. So, the nomologically invariant correlation between the emergent property and the emergence base is a strange and implausible coincidence.

Might it be an option for a strong emergentist to deny that there are nomologically invariant correlations between the microfeatures and the emergent property? That view, together with the plausible assumption that there are worlds in which the physical laws obtain but in which the psychophysical laws do not obtain, implies that zombie twins are not just metaphysically possible but physically possible, that is, there is a physically possible duplicate world in which a person who is physically and functionally indistinguishable from you has no consciousness. That would be a very radical form of dualism. It rules out that consciousness is actually caused by brain states. Of course, there is little agreement on what causation is. However, it is fairly widely accepted that causal relations instantiate laws of physics (see Carroll 1994; Maudlin 2007). So, if our brain states cause consciousness, as strong emergentists tend to hold, then there are strange nomologically invariant correlations between the microfeatures and the emergent property.

Here is a second problem for the two accounts of strong emergence. Both accounts imply that if the emergent property derives from something else, then it derives from something that cannot be found in the natural world. However, as noted above, defenders from both camps deny that the emergent property derives from something else, despite the fact that most accept that it nomologically or metaphysically supervenes on the emergence base. Rather, phenomenal properties are metaphysical primitive or ontologically sui generis (see
O'Connor 2000b). But this assumption is problematic. Both views hold that the emergence base is causally implicated in the *instantiation* of phenomenal properties. But if phenomenal properties are metaphysically primitive, then they do not derive from something more fundamental. So, the universe must have come into existence fully equipped with microphysical properties and ontologically distinct, emergent 'high-level' phenomenal properties. Given that complex life appeared on Earth 9.6 billion years after the Big Bang, these metaphysically primitive high-level phenomenal properties must have existed in an uninstantiated form for more than ten billion years.

On some ways of construing high-level phenomenal properties, this idea is outright incoherent. For example, if the primitive high-level phenomenal properties include properties that organisms now instantiate, then some of them may be properties of representing certain wide contents (e.g., the property of representing *that* thing as red or the property of representing *that* house as being to the right of *me*). But these kinds of properties could not have existed prior to the existence of the external objects on which they depend.

But even if the primitive emergent phenomenal properties do not include properties of representing certain wide contents, the idea of uninstantiated properties is problematic. Presumably some of the primitive emergent phenomenal properties are representational properties (e.g., the phenomenal property of representing redness). However, it is highly implausible to think that abstract entities like properties and propositions, as traditionally construed, can be in the business of representing anything, except in a derivative sense. Traditional accounts of properties and propositions, for example as functions from worlds to extensions, cannot adequately explain how these abstract entities come to have intentional properties that enable them to represent, as there is nothing inherent in these structures that makes them representational. This was the sort of problem that Gottlob Frege (1892a; 1892b; 1918-1919) and the early Bertrand Russell (1903) were struggling to resolve. This has later (somewhat misleadingly) come to be known as 'the problem of the unity of the proposition.' Most proponents of these types of abstract representational entities take their intentional properties to be derivative from certain types of cognitive acts performed by agents. Some define 'propositions' as special types of cognitive acts (see e.g., Soames 2014). But if properties and propositions that can properly be said to have intentionality originate in certain types of cognitive acts, then phenomenal properties cannot be both representational and metaphysically primitive.

Of course, there is the option of saying that it's not the myriad of phenomenal properties attributed to experiences that are metaphysically primitive but the nonrepresentational properties that the rich phenomenal properties have in common: the 'what it's like' of experience. I spell out this type of view below. The problem, though, is that it's not clear that defenders of the two variants of emergentism just outlined can accept this approach. The defenders appear to claim that it's the rich representational phenomenal properties of experiences or a total state of consciousness that count as the metaphysically primitive, emergent properties. But on the alternative approach, it's the nonrepresentational properties that constitute the properties said to be emergent that are metaphysically primitive. This alternative view seems to be a form of panpsychism rather than emergentism proper.

5.3 In Search of Gravitons and Mentons

If strong emergentism is untenable, and standard versions of physicalism cannot account for the nature of consciousness, then it seems that we are left with a version of panpsychism. I am here construing the term 'panpsychism' broadly enough to include what some folks would call 'reductive physicalism.' I elaborate on the reason for this below. For a view to count as a form of panpsychism, as I define the term, it must posit phenomenal properties or protophenomenal properties at the most elementary level of the universe. There are a multiplicity of ways that such a view may go. The most common form of panpsychism is constitutive panpsychism, according to which at least some of the phenomenal properties of macrosubjects are metaphysically determined by phenomenal or protophenomenal properties instantiated by microscopic particles. The most obvious alternative to constitutive panpsychism would be a view that combines panpsychism with strong emergentism; but if strong emergentism is an untenable position, so is a combination of this view with panpsychism. For example, such a view would still need to posit some form of metaphysically primitive phenomenal properties the existence of which would neither fully depend on microfeatures nor on the existence of conscious organisms. Furthermore, it would still leave us with unexplained, invariant correlations between the emergence base and the emergent property. So, it seems to me that for panpsychism to be viable it must be constitutive.

I think that the most plausible version of constitutive panpsychism grows out of thoughts about the analogy between the mystery of consciousness and the mystery of gravity. The mystery of gravity comes from the fact that while physicists can detect the effects of gravity on planets and stars, gravity is very different from other elementary forces. For one thing, it seems to have an indefinite range, permeating matter in the most distant corners of the universe. For another, while we have quite elegant large-scale descriptions of gravity in terms of the curvature of spacetime, we are completely in the dark when it comes to its elementary nature. To compensate for this lack of knowledge, physicists posit the existence of gravitons: hypothetical tiny, massless elementary particles that emanate gravitational fields.

The mystery of consciousness looks not altogether different from that of gravity. Philosophy, psychology, and neuroscience can explain, or may ultimately be able to explain, certain large-scale aspects of consciousness, for example how it arises when neurons oscillate in particular ways. But there is little hope that large-scale theories can offer a theory of the fundamental nature of consciousness.

To be sure, I am not proposing that gravity and consciousness are intimately connected. All the analogy does is demonstrate how two mysteries of science are similar in some interesting respects. Gravity and consciousness are phenomena we are extremely acquainted with in everyday life. Sure, we are acquainted with them in different ways. But both play a crucial role in our everyday activities. Both phenomena allow for some large-scale explanation of how they actually come about. No one has a clue as to the fundamental nature of the phenomena. But the analogy does not end here. For it is at least theoretically possible that both gravity and consciousness are field-like phenomena composed of elementary particles: gravitons and mentons, respectively (see Brogaard 2013a).

To see how consciousness may be a field-like phenomenon, consider John Searle's (2000) distinction between two models of the structure of consciousness: the building block model and the unified field model. According to the building block model, the fundamental units of consciousness are the phenomenal properties of the phenomenal states of an organism. For example, the phenomenal properties associated with my headache at a time are constitutive of my total conscious state at the time. On this view, creature consciousness the property of being conscious instantiated by an organism-derives from the consciousness of a creature's conscious states. On the alternative unified field model, the total conscious state is a unified field that is more fundamental than the individual conscious states. The field changes as representational content is added to or deleted from the conscious field as the brain produces information. 'Information' should here be understood in the standard informationtheoretical sense, according to which a signal has information when receipt of the signal narrows down of the possible ways the world could be relative to what was expected. On the field view, the field itself is not informative but when infused with representations, this results in conscious representations.

There are several good reasons to think that the unified field theory is a more accurate model of consciousness than the building block model. When we search for neural correlates of particular conscious states, what we are searching for typically are neural and glial areas that, if lesioned, negatively affect the quality of the conscious state in question. For example, we take the primary visual cortex to be a neural correlate of visual consciousness because people with lesions to the primary visual cortex either have diminished visual consciousness or lack visual consciousness altogether. But no one of a sane mind would insist on those grounds that the primary visual cortex alone can somehow generate visual consciousness. Feedback loops to the thalamus and the prefrontal cortex, among many other things, are required for an individual to have visual consciousness (see Brogaard 2013b), which suggests that narrow conscious states depend on larger conscious states.

Further evidence for the hypothesis that the total conscious state is more fundamental than any sum of narrow conscious states comes from the fact that it seems that the very seat of consciousness at the level of the brain may be certain regions of the brain stem rather than particular areas of the cerebral cortex. According to Antonio Damasio (see Damasio 2010, 77), two brainstem nuclei, the tractus solitarius and the parabrachial nucleus, which feed into the periaqueductal gray, generate felt states. These felt states in the most modest form are what Damasio calls "primordial feelings." While these primordial feelings cannot account for the phenomenal properties of all conscious states, it is plausible that they are constitutive of all conscious states. That is, the primordial feelings generated by nuclei in the brainstem may be something all conscious states have in common, whereas the differences between our rich conscious states partially originate in information processing in other neural regions. Information processing in the brainstem nuclei is what Francis Crick and Christof Koch call the "enabling factors needed for all forms of consciousness" (Crick and Koch 2003, 119), which they say are not the conscious states that they are primarily interested in. However, the primordial feelings that result from the enabling factors may nonetheless be what a theory of the fundamental nature of consciousness ought to be targeting. If indeed it's these primordial feelings that are at the core of the hard problem, then it's quite plausible that the best model of consciousness is one that treats consciousness as a unified field that can vary in richness and intentionality among organisms and across time.

The view of consciousness that emerges here is a version of constitutive Russellian panpsychism, though there are certain differences between standard constitutive Russellian panpsychism and the view offered here. According to Chalmers (this volume), Russellian panpsychism holds that (i) there are fundamental properties that play the fundamental roles specified in physics, for example there is a fundamental property that plays the mass role, and (ii) some of the fundamental properties are phenomenal or protophenomenal properties. For example, the Higgs boson is the elementary particle that can account for why some fundamental particles have mass. A Russellian panpsychist may say that the intrinsic properties of the Higgs boson are phenomenal properties and that the mass role of the particle originates in these intrinsic phenomenal properties. I take the variant of panpsychism outlined here to be a more plausible variant of Russellian panpsychism than one that takes, say, the intrinsic properties of the Higgs boson to be phenomenal properties. On this alternative Russellian view, elementary particles and forces have fundamental properties that play the fundamental roles specified in physics. However, one of the fundamental forces is primordial consciousness. The properties of the elementary particles that constitute primordial consciousness do not play any of the fundamental roles traditionally specified in physics but they would play one of the fundamental roles specified in a physical theory that takes consciousness seriously, viz. the consciousness role.

If the primordial feeling component of consciousness has explanatory priority, then it's not the intentionality or richness of experience that a theory of the fundamental nature of consciousness needs to explain. Nor is it the subjectivity of experience that needs to be captured by a theory of the fundamental nature of consciousness. Of course, primordial feelings are not *felt* states without subjectivity but a theory of the fundamental nature of consciousness needs to account for the nature of the field that makes the difference between an unconscious creature with subjectivity and a conscious creature.

Evidence from physics tells us that moving fields, or waves, are made up of elementary particles. If we take the analogy with physics at face value, and consciousness is truly like a field, then primordial consciousness presumably is made up of elementary particles. For lack of a better word, call these theoretical constructs 'mentons.' 'Menton' is simply a placeholder for the elementary particle that plays the consciousness role. Mentons are tiny elementary particles with microscopic qualitative properties.

Though mentons and gravitons play completely different roles, we can take the analogy between the particles one step further. The graviton is assumed to be massless because gravity appears to have an unlimited range. Like the hypothetical graviton particle, mentons may turn out to have an unlimited range. Evidence that consciousness has unlimited range comes from the Copenhagen interpretation of quantum mechanics, according to which measurement, or conscious observation, leads to a collapse of quantum entanglement. In quantum entanglement, a pair of particles is generated that has indefinite values with respect to two opposite but correlated properties. For example, an entangled pair may have a two-state spin consisting of spin up and spin down. But there is no answer to the question of which particle has the spin up property and which has the spin down property. When a conscious measurement is made that causes one particle to take on a definite value with respect to the measured property, this also causes the other member to take on a definite value regardless of the distance between them. Moreover, the transfer happens significantly faster than the speed of light, suggesting that the range of some aspects of measurement is unlimited. So, on the Copenhagen interpretation of quantum mechanism, mentons would have unlimited range just like the graviton.

Mentons should not be confused with John Eccles's psychons (see Eccles 1994). Eccles's psychons are theoretical entities that can provide a basis for free will. These hypothetical entities are the psychological components of entities in the cerebral cortex consisting of a psychon and a bundle of dendrites or what Eccles calls a "dendron." Importantly, psychons are not elementary particles; in fact, Eccles rejects panpsychism. They are not microphenomenal qualities constituting phenomenal experience but allegedly identical to macroscopic phenomenal experiences. According to Eccles, it's in the very nature of psychons to bind together to form a unified total state of consciousness, though he does not explain how this binding takes place. Psychons make free will possible, Eccles argues, because psychons corresponding to desires and intentions can interact with dendrons and facilitate neuron firing. His view thus entails that the physical domain is not causally closed. As Eccles's so-called interactional dualism treats phenomenal experiences (the psychons) as a kind of metaphysical primitives, it is a form of emergentism and hence encounters the range of problems outlined above for this type of view. Menton theories avoid these types of problems, because the primitives of this type of theory are microphenomenal qualities that do not depend on the evolution of macrosubjects for their instantiation, and there is furthermore no need to posit fundamental bottom-up psychophysical laws. The fundamental microphysical laws suffice for explaining how microfeatures give rise to macrophenomenal qualities.

Given the menton theory, there is no reason to posit bottom-up psychophysical laws. Physical laws should account for the behavior of all elementary particles. If there are any psychophysical laws, these will obtain at a higher level of complexity. For example, psychophysical laws might help explain how representational properties combine with phenomenal properties to form phenomenal properties that represent specific contents. But unlike the laws posited by the strong emergentist, these laws would not be fundamental laws. They would be derivable, in principle, from physical laws. So, the menton theory avoids the problem of strange invariant correlations between microfeatures and consciousness. Since it does not posit fundamental phenomenal properties that represent anything, the theory also avoids the problem of explaining how fundamental phenomenal properties can represent anything in the absence of psychological subjects, entities in which the very idea of representation is grounded.

It's not implausible that the existence of mentons could be empirically confirmed by physics but that clearly would require a shift in the research strategy of physicists, as they currently are not in the business of searching for the fundamental elements of consciousness. It may be argued that if the existence of mentons could be verified empirically, this would simply show that consciousness is a purely physical phenomenon. This conclusion, however, does not follow. First of all, mentons do not play a role in explaining all purely physical phenomena. For example, they play no role in explaining an electromagnetic field, as this can be explained by appeal to another elementary particle, namely the photon. Further, physical elementary particles do not play a role in explaining the nature of primordial consciousness. This suggests that there is still a meaningful distinction to be drawn between the physical and the phenomenal. Second, an account of consciousness as an entity that's autonomous from other phenomena appears to be assumed by respectable interpretations of quantum mechanics. According to the Copenhagen interpretation, conscious measurements causally determine the values of particles faster than light could travel. The Copenhagen interpretation is far from the only interpretation of quantum mechanics, and it would be beyond the scope of this chapter to outline all the alternatives. Suffice it to say that a majority of alternative interpretations imply that a collapse of the wave function requires conscious measurement. For example, the multiple-universe interpretation assumes that a collapse leads to a branching off of a new universe. In each of the alternative universes conscious measurement causes the entangled pair to assume definite values. On these interpretations, consciousness is autonomous from the underlying physical processes since they exert an irreducible form of downward causal influence. This idea is consistent with the view proposed here but not with reductive physicalism.

Of course, it may be replied that observation and measurement do not require consciousness, and that consciousness therefore need not have any causal effect on the determination of values in quantum entanglement cases. For example, it is plausible that there is a metaphysically possible world in which a zombie—an unconscious physical and functional duplicate of you causes a collapse of the wave function. If this is metaphysically possible, then observation and measurement do not require consciousness, and neither do quantum collapses. However, I think it's important to distinguish between the claim that there *could* be zombie collapses of the wave function and the very different claim that it's not consistent with physical theory that conscious states are ultimately involved in determining the values of entangled pairs. The former claim is plausibly true. But the latter is not given certain interpretations of quantum mechanics. And the latter claim is all we need for it to be true that certain scientific theories treat consciousness as autonomous from physical phenomena. If the elementary menton particles are not truly physical, the proposed version of panpsychism, like most other versions, do not leave the door open to a conceivability argument (aka a zombie argument; see Chalmers 1996). The menton theory does not imply that there cannot be physical and functional duplicates of conscious organisms. It only implies that there cannot be unconscious duplicates of conscious organisms that also are duplicates with respect to all the elementary particles. But zombies—physical and functional duplicates of conscious organisms with no primordial consciousness—are not duplicates with respect to the elementary menton particles. So, even if zombies are conceivable and possible, this would yield no refutation of menton theories.

There is, of course, a rather trivial way in which the menton theory would be conceived of as a form of reductive physicalism. If theories of consciousness that leave it open that physics can empirically confirm the existence of the ultimate constituents of consciousness are a form of reductive physicalism, then this type of theory is a form of reductive physicalism. But on this conception of reductive physicalism, many other forms of panpsychism probably also qualify as reductive physicalism.

5.4 The Combination Problem

Panpsychism and its cousins have long been thought to be plagued by the combination problem (see Seager 1995). William James formulated the problem as follows:

Take a sentence of a dozen words, and take twelve men and tell to each one word. Then stand the men in a row or jam them in a bunch, and let each think of his word as intently as he will; nowhere will there be a consciousness of the whole sentence.... Where the elemental units are supposed to be feelings, the case is in no wise altered. Take a hundred of them, shuffle them and pack them as close together as you can (whatever that might mean); still each remains the same feeling it always was, shut in its own skin, windowless, ignorant of what the other feelings are and mean. There would be a hundred-and-first feeling there, if, when a group or series of such feeling were set up, a consciousness belonging to the group as such should emerge. And this 101st feeling would be a totally new fact; the 100 original feelings might, by a curious physical law, be a signal for its creation, when they came together; but they would have no substantial identity with it, nor it with them, and one could never deduce the one from the others, or (in any intelligible sense) say that they evolved it. (James 1950, 160)

The combination problem challenges theories that posit microexperiences or microsubjects as fundamental to explain how these microexperiences or microsubjects could add up to macroscopic subjective experiences. As far as James is concerned, it is ridiculous to suppose that once we pile up enough microscopic experiences, it all adds up to the familiar macroscopic experience.

James's challenge is taking. However, as Chalmers (this volume) argues, there is, in fact, not a single combination problem but several different ones. Here is one. Only subjects have feelings. Something cannot be felt if there is no subject to feel it. So, on one rendition of the problem James is outlining, the challenge is to account for how myriads of microsubjects can add up to one single experiencing subject.

When the problem is read in this way, it assumes that the microconstituents of macrosubjects are themselves microsubjects with their own experiences and that the macrosubject's new conscious state simply comes together as a result of a particular combination of the microconstituents.

One might argue that the very idea of a microsubject, a tiny microexperiencer embedded within a large-scale subject, is incoherent. Subjectivity, it may be said, cannot reside in things that are not organisms. It's a feature of agents and agency. One subject is not the result of combining several subjects. Mary and Mary-minus, the object that constitutes all of Mary except her right index finger, do not both have subjectivity. Having subjectivity, like being conscious, is an extrinsic, maximal property. This response does not quite solve the problem, however, for microsubjects might just be thought of as the microconstituents that instantiate the microphenomenal qualities.

Following Chalmers (this volume), the subject combination problem can be formulated as follows:

Subject Combination Problem

- 1. If constitutive panpsychism is true, the existence of a number of microsubjects with certain experiences necessitates the existence of a distinct macrosubject.
- 2. It is never the case that the existence of a number of subjects with certain experiences necessitates the existence of a macrosubject.
- 3. Constitutive panpsychism is false.

It's important to bear in mind that a macrosubject here should not be understood as a person consisting of a mind and a body but rather as something like the total conscious state of a person. Otherwise premise 1 is simply trivially false. The existence of a number of microsubjects clearly does not necessitate the existence of a body. Premise 1 appears to follow from the very idea of constitutive panpsychism. The idea is that the total conscious state of the macrosubject supervenes on the microphenomenal qualities of the microsubjects, the bearers of those properties.

Here is an argument for premise 2. The property of being a macrosubject arguably is a maximal property. Ted Sider (2003, 139) defines 'maximal property' as follows: "A property, F, is maximal iff, roughly, large parts of an F are not themselves Fs." As large parts of a macrosubject arguably are not themselves macrosubjects, the property of being a macrosubject is a maximal property. By Lewis's definition (see Lewis 1986, 59–69) of 'intrinsic property,' maximal properties are not intrinsic properties, because whether they are instantiated depends on whether they are part of a larger subject. So, two duplicate entities need not both be macrosubjects. But if the property of being a macrosubject is an extrinsic property, then the sheer existence of a number of subjects with certain experiences does not necessitate the existence of a macrosubject. So, premise 2 is true. Chalmers provides a conceivability argument in support of premise 2, which is reminiscent of James's argument. The first premise of the conceivability argument is the claim that *microsubject S*, *through S*, *exist* does not a priori entail that *microsubject S*₁ *through S*_n *compose a macrosubject*. Assuming that the first premise is genuinely descriptive and hence contains no proper names, it follows that microsubject S₁ through S₂ exist does not metaphysically entail that microsubject S, through S, compose a macrosubject. The conceivability argument is clearly sound. While there is no doubt that the microsubjects compose a mereological sum, no number of microsubjects do by themselves necessitate the existence of a macrosubject with a novel total state of consciousness.

The problematic premise is not 2 but 1. On the type of theory outlined above, primordial consciousness supervenes on the microphenomenal qualities of the elementary particles we called 'mentons.' No conceivability argument can be formulated against this supervenience claim, as it doesn't imply that conscious organisms with only primordial consciousness are possible but only that a certain number of mentons make up a field of primordial consciousness. Premise 1 is false, because most conscious organisms have conscious states that have very rich contents. These rich conscious states do not supervene on the microphenomenal qualities of elementary particles. They supervene on both microphysical and microphenomenal qualities.

Numerous other combination problems arise for theories that solve, or do away with, the subject combination problem. One counterargument that at first glance appears to be particularly problematic for the menton theory is what Chalmers calls the 'palette argument': Here is Chalmers's version (substituting 'a single' for 'a few'):

The Palette Argument

- 1. If constitutive panpsychism is correct, macrophenomenal qualities are constituted by microphenomenal qualities.
- 2. If constitutive panpsychism is correct, there is only a single microphenomenal quality (the menton).
- 3. Macrophenomenal qualities are too diverse to be constituted by a single microphenomenal quality.
- 4. Constitutive panpsychism is false.

Macrophenomenal qualities are phenomenal qualities such as what it is like to see red, what it's like to have a throbbing pain in a tooth, what it's like to feel gloomy, and so on. The single microphenomenal quality posited by menton theories is an elementary particle, a theoretical construct called the Menton. How can a single microphenomenal quality constitute the large range of macrophenomenal properties?

The questionable premise in the argument is premise 3. The single microphenomenal quality does not by itself constitute all the macrophenomenal qualities. Rather, what it constitutes is that primordial feeling that the full range of experiences have in common, the feeling that contributes to there being anything it's like to be in a conscious state in the first place. It's what makes the difference between a subject's experience of red and her zombie twin's analogous physical state. It's consciousness in its most basic and general sense that permeates all conscious experiences, ranging from vague feelings and impulses to sophisticated, rich multisensory experiences.

One of the pieces of evidence that Damasio (2010) uses to argue for primordial feelings grounded in the brainstem is that children with anencephaly, a condition in which the cerebral hemispheres either fail to develop or are seriously compromised by trauma, appear to have a modest degree of consciousness (see Shewmon et al. 1999; Merker 2007). They are awake and exhibit purposive, goal-directed behavior in the form of orienting reactions to environmental events, especially sounds and music. They express pleasure by smiling and laughter, and aversion by fussing, arching of the back and crying, and they seem to have likes and dislikes for things like music. Since these children have no cerebral cortex, they do not have macrophenomenal qualities like phenomenal red or phenomenal green but they likely have macrophenomenal qualities corresponding to likes and dislikes. They presumably also have phenomenal pain qualities but let's set those aside. As affective responses need not be conscious, subjects can express likes and dislikes without any associated consciousness. But if there are unconscious mental states representing likes and dislikes, then simply adding that information to the unified primordial feeling of a subject will generate either a conscious like or a conscious dislike.

While neuroscience can shed no light on the nature of the constituents of the primordial feelings, neuroscience may be able to shed light on how the brain adds qualities (e.g., likes and dislikes) to the unified field of primordial consciousness. Combination, or binding, may happen through gamma oscillations in the brainstem's superior colliculus, which would allow for information processed elsewhere in the brain to gain access to the brainstem nuclei, allowing the information to enter awareness (see Merker 2007).

On this sort of picture, there is little reason to think that more than a single microphenomenal quality is needed to generate the full range of macrophenomenal qualities. So, the third premise in the palette argument is incorrect, and the palette combination problem is avoided.

A third argument against constitutive panpsychism that I will consider here is what Chalmers calls 'the Revelation Argument.' Here is a version:

The Revelation Argument

- 1. The nature of the total state of consciousness is revealed to us in introspection.
- 2. If constitutive panpsychism is correct, the total state of consciousness is constituted by, among other things, microphenomenal qualities.
- 3. Whatever constitutes the total state of consciousness is part of its nature.
- 4. Microphenomenal qualities are not revealed to us in introspection.
- 5. Constitutive panpsychism is false.

The argument certainly does have some initial appeal. Chalmers proposes to reject premise 3. This option becomes plausible once we allow for a notion of nature or essence along the lines of the one proposed by Kit Fine (1994). Fine argues that the standard modal account of essence as de re modality is "fundamentally misguided" (1994, 3). The essential features of an entity are not simply the features that the entity necessarily has. For example, while Kripke's wooden table, Tabby, is necessarily a member of the set {Tabby}, it is not essential to Tabby that it be a member of that set. Nor is it essential to Tabby that seven is prime or that it be such that it's either raining or not. The properties: being a member of the set {Tabby}, being such that seven is prime, and being such that it's either raining or not seem irrelevant to the question of what it is to be Tabby. By contrast, the wood of which Tabby is composed seems relevant to Tabby's essence. Here is an explanation of said intuitions: if there hadn't been sets (or if seven hadn't been prime, ...), then Tabby might still have existed. By contrast, Tabby wouldn't exist if there were no wood. This sort of explanation requires, for its nontriviality and informativeness, that some counterpossibles be nontrivial and informative, or more specifically, that their truth-values be affected by the truth-values of their consequents. Details on how this sort of view may be developed can be found in Brogaard and Salerno (2013).

The important point here is that even if we allow for an account of essence that does not simply define the essential properties of a thing as those it necessarily has, this sort of account would not support a rejection of premise 3. It is typically not the case that if x is constitutive of y, then if there hadn't been x's, then y might still have existed. It certainly is not the case, given panpsychism that if there hadn't been microphenomenal properties, then consciousness might nonetheless still have existed.

So, I don't think we should reject premise 3. I think we should reject premise 1. We can treat revelation as a conjunction of the following two principles (see e.g., Byrne and Hilbert 2007 for analogous principles for color):

Self-Intimation: If it is in the nature of a (rich phenomenal) experience that p, then after careful reflection on the experience it (introspectively) seems to be in the nature of the experience that p

Infallibility: If after careful reflection on an experience it (introspectively) seems to be in the nature of the experience that p, then it is in the nature of the experience that p

It's Self-Intimation that drives the argument: After careful reflection on experience it doesn't seem to be in the nature of experience that it's constituted by microphenomenal qualities (mentons). So, if Self-Intimation is true, then it is not in the nature of experience that it's constituted by microphenomenal qualities. Self-Intimation is not as strong of a thesis as it may initially seem. Many experiences are too fleeting to allow for careful reflection. Suppose in all of your lifetime you only get a very quick glimpse of red. Presumably a quick glimpse of red is not sufficient for careful reflection on your experience of red. But that does not falsify Self-Intimation, of course. The principle, as formulated, is semantically equivalent to:

Self-Intimation: If it is in the nature of S's experience that p *and* S carefully reflects on the experience, then it (introspectively) seems to S to be in the nature of the experience that p

Transient and weak experiences presumably cannot be carefully reflected upon. In those cases, the antecedent is false. So, Self-Intimation is true. The Revelation Argument thus presupposes the premise that there are cases in which we are in a position to carefully reflect on experience, which seems clearly true to me.

However, Self-Intimation is hardly true in general. On an externalist view of perceptual content, it may be in the nature of hallucinatory experiences that they have gappy contents. But careful reflection on a hallucinatory experience does not reveal any gappy contents. It may be argued that while gappy contents constitute hallucinatory experiences, they are not part of the nature of hallucinatory experiences. However, this would imply a rejection of premise 3 in the Revelation Argument. So, the argument fails either way.

Internalists about perceptual content may also deny Self-intimation. Not everything represented by experience need to be discernable even on careful reflection. As some of the sophisticated cases of change blindness reveal, perceivers often do not notice tiny, rapid changes in images despite careful reflection on what they see. But one could hold that the perceptual experience nonetheless represents the changes and that they therefore are part of the nature of the experience, in which case Self-Intimation is false.

One could also offer a speckled hen line of attack on Self-Intimation. If I have an experience that represents a hen with fifty-seven speckles, then regardless of how carefully I reflect on my experience, it may not seem to me that my experience represents a hen with fifty-seven speckles. Even if I could introspectively count the speckles represented by my experience, I might miscount. So, it may seem to me that my experience represents a hen with fiftyeight speckles. But if what experiences represent is part of their nature, then it is part of the nature of the experience that it represents a hen with fifty-seven speckles. Yet even upon careful reflection, it does not seem to me that my experience represents a hen with fifty-seven speckles. So, Self-Intimation is false.

The Revelation Argument, as formulated, concerns rich phenomenal experiences, not primordial consciousness, the what-it's-like feature all rich phenomenal experiences have in common. However, a revelation argument formulated in terms of primordial consciousness appears to be clearly unsound. Here is how it would go:

Primordial Revelation

- 1. If it is in the nature of S's primordial consciousness that p *and* S carefully reflect on primordial consciousness, then it (introspectively) seems to S to be in the nature of the primordial consciousness that p.
- 2. If constitutive panpsychism is correct, primordial consciousness is constituted by microphenomenal qualities.
- 3. Whatever constitutes the total state of consciousness is part of its nature.

- 4. It does not (introspectively) seem to S to be in the nature of primordial consciousness that it's constituted by microphenomenal qualities.
- 5. Either it is not in the nature of S's primordial consciousness that p or S does not carefully reflect on primordial consciousness.
- 6. Constitutive panpsychism is false.

On the reasonable assumption that organisms who can reflect on experience do not have any primordial consciousness that is not integrated into a rich inner life. But given that primordial consciousness is so integrated, it seems clearly false that we can carefully reflect on it. Outside of theory, we cannot 'factor' rich phenomenal properties into a representational component and a (primordial) phenomenal component. So, premise 5 is true, because it is not possible for us to carefully reflect on primordial consciousness. But in that case, the unarticulated premise that S is in a position to reflect on primordial consciousness is false. So, the argument is unsound.

5.5 Conclusion

I have argued that if we take the hard problem seriously, one of the best theories of consciousness is one that takes primordial consciousness, the phenomenal element that all conscious entities have in common, as an elementary force composed of mentons, elementary particles that play the consciousness role. The proposed view is best understood as a kind of unified field theory that takes primordial consciousness to be a field into which informational content can enter and thereby reach awareness. This position fares better than strong emergentism, which is required to posit unexplained correlations between an emergent property and the emergence base as well as intentional properties that do not depend on the existence of conscious organisms. The view also fares better than most other theories that posit microphenomenal qualities at the fundamental level of reality. It does not need to posit that the existence of a certain number of microsubjects necessitates a macrosubject. Furthermore, it can explain how a single microphenomenal experience can lead to a wealth of macrophenomenal experiences.

Notes

^{1.} The principle is based on the principle MS from Ted Sider (2003) and Trenton Merricks (1998). Sider says, "[A]n object O_1 that exemplifies certain intrinsic properties" (Sider 2003, 141) instead of "a living organism O_1 , that exemplifies certain phenomenal properties" and 'compose a living organism O_2 with the same phenomenal properties as O_1 ' instead of 'compose a living organism O_2 with the same phenomenal properties as O_1 .

Bibliography

- Bedau, Mark. "Weak Emergence." Philosophical Perspectives: Mind, Causation, and World. Vol. 11. Ed. James Tomberlin. Malden: Blackwell, 1997. 375–99.
- Broad, Charles. The Mind and Its Place in Nature. London: Routledge and Kegan Paul, 1925.
- Byrne, Alex, and David Hilbert. "Color Primitivism." Erkenntnis 66.1 (2007): 73-105.
- Brogaard, Berit. "The Status of Consciousness in Nature." *The Constitution of Consciousness*. Vol. 2. Ed. Steven Miller. Amsterdam: John Benjamins, 2013a.
- Brogaard, Berit. "Serotonergic Hyperactivity as a Potential Factor in Developmental, Acquired and Drug-Induced Synesthesia." *Frontiers in Neuroscience*. 21. Oct. 2013b. Web. 20 Oct. 2014.
- Brogaard, Berit, and Joe Salerno. "Remarks on Counterpossibles." Synthese 190.4 (2013): 639-60.
- Carroll, John. Laws of Nature. Cambridge: Cambridge University Press, 1994.
- Chalmers, David. "Facing up to the Problem of Consciousness." Journal of Consciousness Studies 2.3 (1995): 200-19.
- Chalmers, David. The Conscious Mind: In Search of a Theory of Conscious Experience. New York: Oxford University Press, 1996.
- Chalmers, David. "Strong and Weak Emergence." *The Re-Emergence of Emergence*. Ed. Philip Clayton and Paul Davies, Oxford: Oxford University Press, 2006. 244–56.
- Chalmers, David. "Panpsychism and Panprotopsychism." *Panpsychism*. Ed. Godehard Brüntrup and Ludwig Jaskolla, Oxford: Oxford University Press, this volume.
- Chalmers, David. "The Combination Problem for Panpsychism." *Panpsychism.* Ed. Godehard Brüntrup and Ludwig Jaskolla, Oxford: Oxford University Press, this volume.
- Crick, Francis, and Christof Koch. "A Framework for Consciousness." *Nature Neuroscience* 6.2 (2003): 119–26.
- Damasio, Antonio. Self Comes to Mind: Constructing the Conscious Brain, New York: Vintage Books, 2010.
- Eccles, John. How the Self Controls Its Brain, Berlin: Springer-Verlag, 1994.
- Fine, Kit. "Essence and Modality." Philosophical Perspectives 8 (1994): 1-16.
- Frege, Gottlob. "Über Sinn und Bedeutung." Zeitschrift fur Philosophie und Philosophische Kritik NF 100 (1892a): 25-50.
- Frege, Gottlob. (1892b) "On Concept and Object." Vierteljahrsschrift für wissenschaftliche Philosophie 16 (1892b): 192–205.
- Frege, Gottlob. "Der Gedanke. Eine logische Untersuchung." Beiträge zur Philosophie des deutschen Idealismus 2 (1918–1919): 58–77. Trans. and rpt. as "The Thought." Mind 65 (1956): 289–311.
- James, William. The Principles of Psychology. Vol. 1. New York: Henry Holt, 1950.
- Kim, Jaegwon. "Making Sense of Emergence." Philosophical Studies 95.1-2 (1999): 3-36.
- Lewis, David. On the Plurality of Worlds. Oxford: Basil Blackwell, 1986.
- Maudlin, Tim. The Metaphysics Within Physics. New York: Oxford University Press, 2007.
- McLaughlin, Brian. "The Rise and Fall of British Emergentism." *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism.* Ed. Ansgar Beckermann, Hans Flohr, and Jaegwon Kim. Berlin: Walter de Gruyter, 1992. 49–93.
- McLaughlin, Brian. "Emergence and Supervenience." Intellectica 25 (1997): 25-43.
- McLaughlin, Brian, and Karen Bennett. "Supervenience." *The Stanford Encyclopedia of Philosophy*, 2011. Web. 21 Oct. 2014.
- Merker, Björn. "Consciousness Without a Cerebral Cortex: A Challenge for Neuroscience and Medicine." *Behavioral and Brain Sciences* 30.1 (2007): 63–81.
- Merricks, Trenton. "Against the Doctrine of Microphysical Supervenience." *Mind* 107.1 (1998): 59–71.
- O'Connor, Timothy. "Emergent Properties." American Philosophical Quarterly 31.2 (1994): 91-104.

- O'Connor, Timothy. "Agency, Mind, and Reductionism." *Persons and Causes*. Ed. Timothy O'Connor. Oxford: Oxford University Press, 2000a. 108–132.
- O'Connor, Timothy. "Causality, Mind, and Free Will." *Philosophical Perspectives* 14 (2000b): 105-117.
- Russell, Bertrand. The Principles of Mathematics. New York: Norton, 1903.
- Seager, William. "Consciousness, Information, and Panpsychism." Journal of Consciousness Studies 2.3 (1995): 272–88.
- Searle, John. "Consciousness." Annual Review of Neuroscience 23 (2000): 557-78.
- Shewmon, Alan, Gregory Holmes, and Paul Byrne. "Consciousness in Congenitally Decorticate Children: Developmental Vegetative State as Self-Fulfilling Prophecy." Developmental Medicine and Child Neurology 41.6 (1999): 364–74.
- Sider, Ted. "Maximality and Microphysical Supervenience." *Philosophy and Phenomenological Research* 66.1 (2003): 139–49.
- Soames, Scott. "Cognitive Propositions." *New Thinking about Propositions*. Ed. Jeffrey King, Scott Soames, and Jeff Speaks. New York: Oxford University Press, 2014. 91–126.
- Van Cleve, James. "Mind-Dust or Magic? Panpsychism Versus Emergence." Philosophical Perspectives 4 (1990): 215–26.
- Van Gulick, Robert. "Reduction, Emergence and Other Recent Options on the Mind/ Body Problem: A Philosophical Overview." *Journal of Consciousness Studies* 8.9–10 (2001): 1–34.

6

Land Ho? We Are Close to a Synoptic Understanding of Consciousness

GREGG ROSENBERG

6.1 Introduction

In 2004's *A Place for Consciousness (APFC)*, I proposed a new metaphysics called the Theory of Natural Individuals (TNI), which was a reimagining of causality. I demonstrated how the new metaphysics not only provides new answers to issues at the foundations of causality and philosophy of nature, but also can help us understand why consciousness exists. In its preface, *APFC* set itself the goal of answering four fundamental metaphysical mysteries surrounding consciousness:

- 1. Why should the intrinsic properties of a physical system ever be experiential?
- 2. Why do experiential properties exist above the level of the microphysical, enabling large scale cognitive systems to experience macrolevel intrinsic content?
- 3. Why should experiential qualities form a unity of the kind we are acquainted with in consciousness?

and

4. Why should phenomenal character, as the intrinsic content of the physical, correspond so closely to the information structure within the brain?

In this chapter I am going to provide an overview of TNI and place it in context with recent advances in our scientific understanding, with the goal of describing how we are close now to having a synoptic understanding of consciousness from the most general level of metaphysical questioning to the most concrete level of biological understanding. What this means is not that all detailed questions are close to being answered but rather, very importantly, the emerging set of paradigms are close to moving us from the 'Big Mystery' phase of understanding consciousness to the 'Puzzle Solving' phase at every level of questioning. In sections 6.2 through 6.5, I am going to review the Theory of Natural Individuals introduced by *APFC*. In sections 6.6 and 6.7, I am going to introduce and discuss the Synoptic Pyramid for understanding consciousness, which puts TNI into context with the emerging science to paint a picture of complementary and coherent answers now available to the big questions at multiple levels of inquiry.

6.2 Reframing Which Problems Matter

Take a sentence of a dozen words, and take twelve men and tell to each one word. Then stand the men in a row or jam them in a bunch, and let each think of his word as intently as he will; nowhere will there be a consciousness of the whole sentence.... Where the elemental units are supposed to be feelings, the case is in no wise altered. Take a hundred of them, shuffle them and pack them as close together as you can (whatever that might mean); still each remains the same feeling it always was, shut in its own skin, windowless, ignorant of what the other feelings are and mean. There would be a hundred-and-first feeling there, if, when a group or series of such feeling were set up, a consciousness belonging to the group as such should emerge. And this 101st feeling would be a totally new fact; the 100 original feelings might, by a curious physical law, be a signal for its *creation*, when they came together; but they would have no substantial identity with it, nor it with them, and one could never deduce the one from the others, or (in any intelligible sense) say that they evolved it. (James 1890/1950, 160)

With this well-known quote, James introduced a couple of difficulties that we have come to call 'the combination problem' for panpsychists. The first is alluded to by the image of that poor twelve-word sentence that will never exist because its component words are trapped in the heads of twelve men. This is a *composition of feelings* problem. What rules of composition could blend separate feelings (or thoughts) into a single feeling or thought?

The second image, of the hundred "windowless" feelings, imagines success overcoming the first problem but suggests the cost of success is a new *composition of experiencers* problem. The only way to get a new feeling more complex than the originals is to create a new substantial identity to feel it, and this new substantial identity will be strongly emergent from the feelings which occasioned its emergence. This is a problem because panpsychism is supposed to be an alternative to emergentism.

James was substantially right that panpsychism leads to strong emergentism. However, he and others have been wrong to frame panpsychism and strong emergentism as competing alternatives. Both panpsychism and strong emergentism are implied by a much deeper theory, which is attractive because of the way it makes sense of the metaphysics of causality, and many other questions in metaphysics, as well as providing insight into important questions in the philosophy of mind. Panpsychism shouldn't be adopted simply because it might be a solution to the mind-body problem. It is also a solution to the carrier-causality problem, the problem of what intrinsic natures carry the schemas of causality in our world. This makes a large difference in how many things related to panpsychism should be viewed, particularly the motivations for adopting it, its explanatory power and its relation to emergence.

To reach this point of view, one has to switch to a new frame on certain problems, so that oppositions and concerning questions that recur in many contemporary discussions of panpychism and/or the mind-body problem seem less relevant. It has been years and years since I have cared whether a view is dualist, monist, or pluralist; or whether something was strongly emergent or weakly emergent; whether a view was panpsychist or merely neutral monist; whether causes have to precede effects or why there is something rather than nothing. The reason is because TNI provides a different frame on philosophy of nature, in which different things seem important.

What follows is a discussion of a series of frame shifts that change one's point of view on whether certain ontological outcomes are 'good' or 'bad' and whether certain questions are interesting or not.

6.3 Changing the First Frame: From a Combination Problem to a Boundary Problem

James asks us to imagine twelve people thinking of twelve words. Crowd them together and we do not create a sentence. This seems like a hard problem to resolve. But is it any harder than the following problems: Imagine eight glasses of water. Crowd them together as close as you like, and they do not make a puddle. Imagine one hundred organs in a freezer. Make the freezer as small and cramped as you like, but they do not make an animal. Imagine a billion cells in a petri dish. Make the petri dish as crowded as you like, they do not form an organ.

There are related problems, which are not hard: Imagine twelve waves flowing toward a beach. Have them meet, and the twelve waves combine into a single wave. Imagine two electric fields combined to create a more powerful electric field. Imagine a billion cells in a fetus, combining to make an organ.

This second set of related problems show there are good combination rules for many different kinds of things, in which these things can combine to become another unitary thing of either the same kind or a different kind. This is not a *conceptual* problem. It's a problem about understanding nature deeply enough to know when things interact and combine, and when they don't. In some cases, the combination rules are matrix or vector addition rules for things like waves and fields, in which different things can combine to make new things and the old things are lost. In other cases, the combination rules are rules of tight coupling and interaction, in which the old things combine to make qualitatively new things while still maintaining an integrity of their own.

We can imagine a panpsychist world in which phenomenal characters combine via matrix or vector-like addition rules, while experiencers combine via tight interaction rules. The difficulty is really the odd empirical facts about bounded phenomenal fields existing surprisingly at a midlevel of the physical world, at a scale corresponding to physical activity in animal brains.

A panpsychist might predict primordial experiencers at the base of the world, simple fireflies of flickering experiencers at low levels of space time, without any experiencers at higher levels: feeling, feeling everywhere, but not a drop can think.

A panpsychist can avoid this surreal picture via appeal to known types of combination rules. In a world where experiences combined according to matrix or vector rules, for example, one might predict an ocean of experience constituting universal consciousness, but no experiencers below that level: a comically cosmic god-mind, experiencing itself alone forever.

The Scylla and Charybdis of these two alternatives are what I call the Boundary Problem for Phenomenal Individuals. How can one allow experiences to combine from the low level to the midlevel but in a structured way that does not run away from us? It is a reframing of the combination problem in terms of a different but related problem.

The hard nut of the Boundary Problem is that animal experiencers possess a kind of inherent individuality at a physical midlevel of reality, which is hard to explain. If panpsychism is true, why do the boundaries exist *just so*? Panpsychists can easily find types of combination rules, which could work to overcome the combination problem. However, once a panpsychist recognizes a known kind of combination rule might be in play, he or she can coherently hypothesize almost as many ways of determining boundaries for experiencing subjects as there are of seeing organization in the world's pattern of microphysical interaction. Constraining application of the combination rules to produce a meaningful outcome on the Boundary Problem is hard. Moral: Boundaries are harder to explain than combination. We are faced with the need to understand more deeply what it is to be an inherent individual in the natural world.

There is one relation at the borderline of physics and metaphysics, which divides the world by its very nature, the relation of causal interaction. Causal interactions imply partitions: they divide the world into different, mutually influencing spaces, and do so at many levels of nature. We might discover, by looking at causality and causal interaction, that causal interactions have certain important aspects that distinguish natural individuals. For panpsychists, these individuals would be attractive candidates as the supporting ontology of experiencing subjects.

6.4 Changing the Second Frame: From Cause and Effect to Causal Significance

Mainstream discussion of causality almost always begins and ends with discussions about judgments of cause and effect or the language of cause and effect. The guiding image is that of a probability tree. Time is a tree of different outcomes, the actual world is a path through the tree, and causality is a way of influencing or setting probabilities for each branch of the tree.

A 'probability tree' in which causality is a way of setting/changing probabilities is the fundamental frame from which discussions of causality depart. Significant questions within this frame include is causality epistemic or metaphysical? Can causality work only from the past to the future? How do we understand the relation between 'cause' and 'effect'? What are dispositions? This fundamental causality frame does not help a Russellian monist or panpsychist like me, whose interest in causality is driven by a desire to understand causal interactions and the way causal interactions partition the world.

The probability tree is a flawed metaphor. The major problem with the probability tree is that the cause-and-effect relationship is a poor starting point for understanding the metaphysical basis of interactions and its relation to the world's structure, as our cause-and-effect concepts don't get at the essence of causal connection (the causal nexus), and contain many assumptions that could be false in a world where things interact. It is also heavily intertwined with human psychology and perception.

For example, dynamics are necessary for cause and effect but not necessary for interesting causality in general. If different parts of the world constrained one another so that counterfactuals about states of the world as a whole were true, one could say interesting things about laws and about causal constraints between different parts of the world, even without dynamics. Concretely, imagine a timeless world, which was nothing but an unchanging Rubik's Cube. Some combination of colors on its six faces is showing. Other combinations *could* show, and there is an atemporal physics describing which ones are possible and which ones are not (there is no physics of transitions). There are meaningful causal constraints in terms of constraints between components in the Rubik's Cube world, even though there is no cause and effect.

Of course examples of the subjective nature of cause and effect are wellknown. The idea of 'cause' is closely related to agency and blame, and 'effect' to goals and means. Negative facts such as absences can be said to cause things. Figure-ground relations, levels of granularity and so forth all impact causeand-effect judgments. And just try to get agreement about cause and effect in a political situation, such as the Israeli-Palestinian conflict.

Our common concept of cause and effect also builds in several other parameters: locality, directionality, categorical constraints on relata, the arity of the relation, all of which are baggage on a more basic notion. To think of causality in a fresh way, we need to get rid of the baggage and pare it down to its core essential truth: The world is in a constrained state, in that having one part of the world in one state places constraints on what states other parts of the world can be in. This truth captures the fundamental mystery of non-Humean causation.

It doesn't matter if the parts of the world are in different time slices; or if they are local; or if there are two of them versus three, four, or an infinite number. It doesn't matter if we think of these parts of the world as events or states or facts. It doesn't matter if one is a 'cause' and the other is an 'effect.' If one part of the world being in one state, places a meaningful physical or metaphysical constraint on the state of another part of the world, we are presented with a mystery of causality.

To express this mystery, I like to talk about the Two Canvases of Causation.

Imagine two great, blank canvases that you paint with color one drop at a time. Imagine also that the two canvases are two very different kinds of surfaces to work with. You call the first canvas the Humean Canvas. It acts like a normal canvas, as it will accept any drop of paint anywhere on its surface in any color that falls on it. If you let a drop of red paint fall onto the Humean canvas, it will stick where it lands. The same will happen if you then drop a speck of yellow paint somewhere else on the canvas. You can fill the whole canvas this way, with whatever colors you want, anywhere you want.

You call the second canvas the Canvas of Causation. It is a marvel. If your first drop of paint is a bit of green, and then you try to place a drop of red next to it, the red paint will bounce off. The canvas will not accept it. But it will accept yellow. And the more paint you put on the canvas, the pickier it seems to become. Each bit of color you put on the surface seems to place a constraint on what colors may appear elsewhere on its surface. Although the canvas will

allow you to paint it in many different ways, it will accept only combinations of color that make for a beautifully covered canvas. If you try to pour many colors on it all at the same time, some will stick and some will run off, and each time what remains after the run-off will be a beautiful pattern of color. Every color and every drop matters, jointly enforcing or excluding the colors that will finally appear on the canvas.

Although the Humean Canvas is ordinary, the Canvas of Causation seems like magic. Yet the Canvas of Causation seems to be more like the world we actually live in. It is a world in which nature includes and excludes membership based on what else has made it into the club. Making our world be like a Canvas of Causation requires some extra ingredient over and above simply having a world in which things can happen. This extra ingredient, whatever it is, should be what a theory of causality is about.

The central concept for causality, in this view, is *causal significance*, not cause and effect. The *causal significance* of a thing is the constraint its existence adds to the space of naturally possible ways the world can be. Though it covered other ground, *A Place for Consciousness* was chiefly about introducing and describing the metaphysics of causal significance and demonstrating its explanatory power.

An emphasis on *causal significance* is a second kind of reframing that Russellian monists and panpsychists need to make progress on their world view. They need to reframe the problem of causality from being a problem about cause and effect to being a problem about causal significance, because a theory of causal significance yields many unexpected and very important explanatory benefits, not only in the philosophy of mind but elsewhere.

The shift away from cause and effect as the core concept of causality also requires shifting away from probability trees as the core framing device for discussions of causality. If we emphasize causal significance, it seems the central job of causation is to select potentialities in ways which constrain possible states of the world. In other words, causality doesn't just make *future probabilities*. It makes the *world actual*. The proper framing device for causal significance is to think of causality as a *potentiality filter*, rather than a probability tree. On one side of the filter is a set of potentialities, and on the other side is a smaller set.¹

As soon as one starts thinking of causality using potentiality filters as a tool, it just becomes obvious that there is nothing inherently temporal about a potentiality filter. The potentialities fed into a potentiality filter don't need to flow from past states to future states of individuals. Efforts to restrict the formalism that way look obviously artificial and hard to justify.

The potentialities fed to the potentiality filter could come from individuals at a lower level of nature and be fed into an individual at a higher level. This scenario is mathematically, logically and metaphysically coherent. It also coheres well with known physics.

Indeed, it becomes equally obvious that a core, unargued assumption in much of classical metaphysics is that things at lower levels of nature are already, in and of themselves, in determinate states, just as we find things at higher levels in ordinary experience. It is also assumed that things at higher levels of nature *must* inherit their determinate states from the determinate states of the lower-level things, which compose them. In philosophical parlance, classical metaphysics assumes that determinate macrolevel states of the world *must* strongly supervene on independently determinate microlevel states of the world. This statement is a kind of *microdetermination thesis*, or *mic-d* for short.

Why do we believe *mic-d*? There is, in fact, no evidence for it and quite a large amount of evidence against it.

The only real reason for believing it, is it is 'intuitive.' But that is no reason at all. We have had no conceptual tools for thinking of anything different and so it is 'intuitive' by default. Additionally, we know enough now to say with confidence that questions about fundamental nature tend to not have intuitive answers, and that things at the lowest levels of nature are not in and of themselves in determinate states.

The problem is, panpsychists (and others interested) haven't had the vocabulary to raise this as an issue or a formalism to model how an alternative could work or be resolved. The assumption has been buried so deep, there's been no good way to talk about it. There also has been no good way to model worlds in which it is true and worlds in which it is false, to examine the differences between those worlds, and the implications of those differences. To get anywhere, we need to shift from modeling probability trees extended through time, to modeling ladders of potentiality extended through levels of nature. A significant part of my project in *A Place for Consciousness* was to provide a vocabulary and formal model for asking these questions, representing differing answers, and exploring implications of those representations. What I found was that model worlds in which *mic-d* is false

- i. tend to have physics that make the physics of our world with its randomness, backward causality, indeterminate states, nonlocal causality and measurement problem, look expected rather than weird;
- ii. deductively require, just to make causality work, the presence of subjectivity with many of the odd properties consciousness seems to have (e.g., unified fields of intrinsic properties structured according to information-theoretic constraints);

- iii. possess higher-level individuals that can easily extend to the midlevels of the natural world;
- iv. recast traditional oppositions between things like emergence and panpsychism, or dualism and monism, or mental causation and epiphenomenalism in ways which take the bite out of them.

6.5 The Theory of Causal Significance

In *A Place for Consciousness*, I introduced a vocabulary and formalism for representing both the truth and falsity of mic-d within a single metaphysics for causality, and for exploring consequences of a metaphysics in which mic-d could actually be an open question. The result of this effort was the Theory of Natural Individuals (TNI), in which the Theory of Causal Significance (TCS) was its heart.

TCS describes the causal connection as a potentiality filter, of a very general nature, in which the same metaphysical device produces temporal connections of cause and effect, causal interactions between individuals within a time slice, and also higher-level individuals that can make indeterminate lower levels more determinate. TCS is a unifying framework, in which these three things are all aspects of the same thing, looked at from different angles.

By achieving this unifying effect, TCS provides us a precise way to think of a world of layered individuals, in which determinate individuals can exist on top of 'indeterminate' lower-level individuals and make them determinate in turn. It does this by giving us a precise way to think of what an indeterminate individual is, and how it might become determinate: "This kind of indefinite disjunctive state becomes easier to grasp once one realizes that a disjunctive state such as *s1 or s2* is logically equivalent to the conjunctive state: *potentially s1 and potentially s2 and not potentially anything else.* The indeterminate state of an individual is equivalent to a definite (note: *determinate*) state of that individual understood as a pluralistic selection from its space of potentialities" (Rosenberg 2004, 209).

What TCS embraces is a situation where the answer to whether an individual is in a determinate state or an indeterminate state depends on the perspective of the description. An individual may be said to be in an *indeterminate* state from the perspective that asks which of its potentials would be experienced (directly or via its influence on other individuals). TCS equates the experienced potentials of an individual with its realized or actual state and allows that the state of some individuals may be indeterminate, considered independently of their causal relations within larger individuals. But that same individual can be said to be in a *determinate* state from the perspective that asks what subset of its potentials could be experienced (i.e., realized/actual). Additionally, TCS implies a realism about potentiality, in which unrealized potentialities can have causal significance. That is, the experienced/actual state of the world can depend on states of the world that could have been experienced but aren't.

This feature of TCS provides an important anticipation of the general form of quantum physics. It is a model of the causal nexus that provides a natural metaphysics for nonepistemic interpretations of the wave function in physics, in which each eigenstate is considered a real potential, with real causal power, but there is indeterminacy regarding which potential an observer would encounter. TCS provides a metaphysics in which there is an equivalency between saying an individual is in an indeterminate state (*s1 or s2*) with regard to what could be its actual observed state, and saying the individual is in a determinate state (*potentially s1 and potentially s2 and not potentially anything else*) with regard to its real causal significance.

By adopting TCS, we can ask metaphysical questions that previously were hard to ask, and we can provide answers that previously were hard to even represent. Importantly for the panpsychist, TCS provides a solution to the combination problem that does not fall foul of the boundary problem.

TCS has a simple ontology, consisting of (i) two kinds of properties, (ii) one relation that holds between them, and (iii) intrinsic causal laws. Here is the furniture of TCS:

Property Types

- *Effective properties*—Properties that intrinsically/inherently contribute to constraints on a causal nexus.
- *Receptive properties*—Connective properties that intrinsically/inherently bind to effective properties, creating a causal nexus.

Fundamental Relation

- *Binding*—A primitive metaphysical relation in which one property becomes part of the completed essence of another property.
- *Causal laws*—Laws describing relationships of compatibility, incompatibility, and requirement between effective properties.

The ideology of TCS is slightly more complex than the ontology, as it introduces two new concepts: receptive properties, and the concept of *completion*, which occurs in the definition of the binding relationship. Receptive properties in TCS play the role of the causal connection. They are the potentiality filters existing between effective individuals allowing them to constrain one another's potential states. It's important that there is a receptive intermediary in the causal nexus: if binding occurred directly between effective properties, the model could not simultaneously represent their interactions and establish boundaries and a plurality of individuals, some at different levels of nature than others. The world would necessarily be a single-level, homogenous mesh.

In TCS, two or more effective properties can bind to the same receptive connection, which acts as a context where the intrinsic constraints between them can be activated and structured. Think of receptivity as a kind of neutral background field, whose essence contains the possibility of interaction between effective properties and that establishes the structure of interaction and causality in the model.

The concept of *completion* is an important ideological innovation in the theory. It is part of a cluster of concepts that also include *determination* and *concreteness*. The idea is that properties considered in isolation from their binding relation to other properties do not have *determinate* states. Lack of determinateness is interpreted to indicate a kind of essential incompleteness in a property's or individual's isolated nature. Properties *need* context (causal binding to other things) to be anything in particular.

In TCS, the state of such a context-free property gets represented as a disjunction of states (or, equivalently, a conjunction of potential states). Recall, this is the meaning of an individual being in an *indeterminate* state. When TCS says a causal relation makes an individual more determinate, it means there has been a reduction in the number of disjuncts needed to represent its state (or number of conjuncts in its list of potential states). It is indeterminacy about the character of causal influence that could be experienced by individuals on the other side of the causal connection. One can think of this on analogy to eigenstates going flat during a process of decoherence in an object's wave function or collapse in a collapse interpretation. TCS says such indeterminate properties considered in isolation from their contexts are incomplete, and they need to bind to other properties to (take on context and) become complete. This process of binding and completion is what causal interactions *are*.

The model example for thinking about this is to consider a property like electron spin. Inherently, electron spin can be either spin up or spin down. Considered in isolation from the context of an individual electron in the world—that is, considered as an essence—electron spin as a property has an indeterminate state in TCS, which is represented as a disjunction of up and down values. According to TCS, spin, considered in isolation, is therefore *abstract*. Spin is an example of an effective property. An instance of electron spin starts to become concrete when the abstract spin binds to a receptive property (the abstract may bind to many receptive properties, each of which becomes a different individual). A basic individual, such as an electron, is modeled as multiple effective properties (e.g., spin, charge, mass) bound to a common receptivity. The common receptivity gives that individual its causal unity as a particle. If the combination of properties is not further constrained, it will inherit the indeterminateness of its spin, so the particle itself may need to bind to become determinate. This process may repeat again and again, until the constraint structure is strong enough to make spin take on a determinate value and the particle be a determinate type.

The metaphysics here is one in which effective properties are determinable universals, and causal relations (in the form of receptive connections) are relations with individualized intrinsic essences that bind to universals. In binding to them, they provide context in which only some of the universal's potentials can be individualized as part of the concrete world. Causal relations are therefore the engine of creation, which take the world from abstract to concrete in a stepwise fashion, through layers of receptive binding, and the vector that represents the universal's loss of potentiality through layers of these relations was called an *ingression* in *APFC*. Each ingression was an event, and events could be bound together into processes. The ontology is therefore an event, not a substance, ontology, and there are no enduring characters in the concrete world.

This ideology is exotic compared to classical metaphysics. In the most relevant sense, I argue this exoticness is not a drawback, as it is clearly a metaphysical recasting of well-evidenced aspects of the standard model of modern physics. One might say it scores low on *intuitiveness* (it is not the first ideology one might think of), but I argue in *APFC* it is high on *plausibility* (it is by far the ideology most coherent with our actual evidence about causality). I believe TCS is a simple theory ontologically, a moderately complex theory ideologically, and I believe a very plausible theory from the standpoint of evidence, of what we actually have reason to believe about nature.

TCS is the heart of a theory of natural individuals because it contains at its basis a recursive combination rule, which describes how causality creates natural individuals and can do so above the base levels of reality. The combination rule is simple:

- Base rule: Any primitive receptive or primitive effective property is a natural individual.
- Recursive combination rule: Any receptive property, which completes itself by binding to two or more other natural individuals is a natural individual.

The combination of effective properties that occurs within a binding relation is treated like a matrix or vector combination of properties, occurring within the boundary provided by the receptive connection. The receptive property itself is irreducible causality, an emergent.

The purpose of a receptive connection—its occasion for existing and what it produces—is a determinate state, or a *more* determinate state, for otherwise indeterminate individuals. Receptive connections as emergents are not otiose or indulgent. They are metaphysical ground zero. They have an essential function in causality as the engine of creation, which must somehow make a determinate world from a set of indeterminate and abstract potentials. Completion happens because a receptive context enables additional constraints to be active on indeterminate individuals, which make them more determinate.

TCS thus gives a formalized account of causal significance by describing how a potentiality filter—called a receptive connection—can metaphysically enable causality in the world by producing determinate states of individuals from relations between indeterminate properties. Once formalized, the theory allows us to model causality in a world that is intrinsically determinate at its lowest level, as in classical physics (though this requires some ugly maneuvering, that speaks against the likelihood there could be a world with classical physics); but it also allows us to model worlds that are inherently *indeterminate at the lowest level but made determinate by the presence of higher-level individuals*. Additionally, it shows how the same mechanics at work between levels can produce dynamic causality across time slices.

Of the many kinds of worlds that can be modeled using TCS, the nonclassical, multilevel worlds with inherently indeterminate lower-level individuals (in *APFC*, called 'indeterminate when considered independently') are by far the most natural. Also, the intralevel structures that model interactions between individuals and the interlevel composition relations between levels of nature are rich and provide more insight into important metaphysical questions than does an investigation of cause-and-effect relations across time.

For example, one can clearly show a causal role for higher-level individuals (by implication, animal consciousness) that is different than anything ever pictured in philosophical discussions of interactionism or emergent causation. The role of higher-level individuals is to act more like final causes, because the determination of the state of the higher-level individual is a cause of determination at the lower levels, without imposing any interaction force across levels or impacting energy. It is simply a selection force. In the reverse direction, the relation between the lower-level individuals and the higher-level individuals is more like material causation rather than pure composition (because of the presence of irreducible causal connection in the form of receptive connections). This model then provides a potential third way to think of mental causation, aside from interactionism and epiphenomenalism, in a nonphysicalist metaphysics. Mental states are just states of certain higher-level individuals, and they are part of a universal process of causality in which there are twoway causal relations between levels of nature. Higher levels provide contexts for constraint, which make them final causes for the determination of otherwise indeterminate base levels; and the base levels, made determinate, provide material causation for the higher levels. Effective causation remains limited to within-level relations.

There is a kind of strong emergentism in TCS, but it is a kind that renders untroubling the usual discussions and concerns about strong emergentism. In TCS, receptive connections bind individuals at different levels of nature, and they are not reducible. Each receptive connection is a constitutive property of its own individual, not composed from lower-level properties. But recall, the receptive connection is just *causality*. It is happening just the same everywhere—it binds fundamental properties into particles and particles into interacting systems—and is not a special thing brought in suddenly at one level of nature, or just to explain mental properties. It is in no way ad hoc or surprising.

Other traditional discussions, such as the ontological 'counting' discussion, also become less interesting. One could call this a dualism of receptive and effective properties. Or a monism of the causal nexus. Or even a pluralism of intrinsic and extrinsic aspects of causality. It is a naturalism of the causal mesh or a kind of property Platonism of unrealized abstract potentials. One could claim this is a kind of Russellian neutral monism, a kind of physicalism, or a kind of nonphysicalism. However, from the perspective of TCS, this sort of counting question does not advance any issue and does not grip the imagination. What matters is how much it explains—which is a lot—relative to how much it assumes, which is not so much.

6.6 A Synoptic Understanding of Consciousness

TCS, for all its virtues, does not avoid the kinds of critiques of physical theory put forward by Russell and Whitehead. The natural individuals in TCS need intrinsic properties to perform the functional roles laid out by the theory. These intrinsic properties are called carriers, as they carry the extrinsic descriptions for effective and receptive properties within the theory. Adding this postulate to TCS yields the Carrier Theory of Causation (CTC) and together the two theories, the Theory of Causal Significance and the Carrier Theory of Causation, make up the Theory of Natural Individuals (TNI).

The Carrier Theory of Causation is an additional postulate to TCS, but because TNI contains both theories as components, we are able to use TNI to deduce specific requirements about what the intrinsic carriers of effective properties must be like. One can deduce, for example, that the carriers of effective properties would have to possess a kind of unity similar to the unity we find in consciousness and would have to have intrinsic relations of compatibility and incompatibility similar to what we find among phenomenal properties like red and green. Also, using TNI we can deduce that the intrinsic carriers of effective properties must support scalar relations also similar to what we see among phenomenal properties in their intensity dimensions (e.g., the loudness of a tone).

When we turn our attentions to receptive carriers, we can deduce the carrier of receptivity must have a kind of contentless openness, similar to what meditative practices report as characteristic of the pure experiencing self, and that the receptive carrier's relationship to effective properties would have to be much like what we see as the relationship between experiencing and phenomenal properties.

For these and other reasons, someone trying to make sense of TCS in the world will feel compelled to hypothesize that phenomenal properties are the intrinsic basis of effective properties; and that an experiential property is the intrinsic basis of receptive connection; and that the causal nexus in our world is carried by the experiencing of phenomenal properties by the carrier of the receptive connection. The experiencing of phenomenal properties *is* the causal nexus in our world. Anywhere there is direct interaction between natural individuals, there we will find the occurrence of experiencing. So TNI endorses a panexperientialist version of Russellian monism, taking much inspiration from Whitehead in its specific form.

TNI is a metaphysics, and it is interesting to consider TNI in the context of emerging scientific candidates for explaining consciousness. Over the last few decades we have seen the formulation, one by one, of several very intriguing and interesting candidates for a scientific explanation of consciousness: Bernie Baars Global Workspace Theory (see Baars 1988; Baars 2002); Dynamic Core and Thalamo-Cortical Resonance models as developed by Edelman, Tononi, and others (see Edelman and Tononi 2000); and most recently the Integrated Information View being developed by Tononi (see Tononi 2012) and the growing community inspired by his original work. Below I will introduce a very interesting frame I call the Synoptic Pyramid that places these views together with TNI, ordering them from the most general at the base of the pyramid to the least general at the top.



Figure 6.1 The Synoptic Pyramid.

Reading the pyramid from the top down, three things stand out:

- 1. The journey from pyramid top to pyramid foundation represents a consistent generalization and universality of theory based on core insights. Each higher theory is able to function as a bridge from the realm of the theory below it to a more specific topic domain. And lower theory can in important respects be seen as anticipating theory above it.
- 2. There is a common theme of connection and integration running through all theories.
 - a. At the foundation: TNI is explicitly about a nexus of elements mutually enforcing state compatibility by passing constraints to one another through an informational signaling system.
 - b. At level two: IIT takes the form of a quantity describing a system of mutual constraint among many elements via the integration of shared information.
 - c. At level three: Global Workspace Theory is about a computational system enforcing a constraint of semantic consistency and integrated control among elements in a global workspace.
 - d. At the top: The Dynamic Core/Thalamo-Cortical Resonance theory describes a system of densely connected physical feedback loops between neural systems, creating coherent behavior between separated brain regions.

3. From the bottom to the top of the pyramid, we can see each higher theory as an interesting direction for exploring implementation of the ideas in the theory below it. Together, they are highly coherent with one another and point to a consistent, synoptic world view for fitting consciousness into nature with mutual causal connection at the core.

There's a synoptic picture implicit here, in the sense of a broad and integrated paradigm for understanding consciousness at multiple levels of inquiry. First consider the four questions *APFC* set for itself and their answers (from *APFC* chapters 12–14),

1. Why should the intrinsic properties of a physical system ever be experiential?

Answer: Because subjective experience has the properties metaphysically required of an intrinsic carrier of mutual causal constraint within a nexus of causality: (i) Phenomenal character has the properties necessary to carry effective constraints; (ii) Subjects of experience have the properties necessary to carry receptive connectivity;(iii) binding relationships between subjects and phenomenal character in experiencing have the metaphysical intimacy required to carry causality in the causal nexus.

2. Why do experiential properties exist above the level of the microphysical, enabling large-scale cognitive systems to experience macrolevel intrinsic content?

Answer: Because individuals at lower levels are not always in determinate states, independently of the existence of higher-level individuals, so the existence of determinate states necessarily implies the existence of higher-level causal nexuses that make the world determinate.

3. Why should experiential qualities form a unity of the kind we are acquainted with in consciousness?

Answer: Because experiential qualities carry effective constraints, which are holistic, informational properties that can only exist within the larger context of a causal nexus.

4. Why should phenomenal character, as the intrinsic content of the physical, correspond so closely to the information structure within the brain?

Answer: Because phenomenal characters, as carriers of effective properties, play the role of signals carrying information about constraints within the nexus, and the structure of the signals necessarily reflects the information structure of elements within the nexus rather than their physical structure. However, TNI stops without providing concrete answers to questions about how to measure whether an experiencing causal nexus is present. *APFC* does describe some conditions on an answer. It predicts that an appropriate theory will have panexperientialist consequences and it puts a high-level constraint on what such a theory should look like:

A completed receptive connection has (1) at least one constituent with an indeterminate state when considered independently of its membership in the nexus, and (2) a common receptivity being shared by two or more constituents. The shared receptivity establishes a connection between the members of the nexus through which they contribute to a set of simultaneous constraints on their joint states.

Associated with each natural individual I_k is some set of rules, label it Λ , such that, the set of constraints in Λ is most naturally thought of as containing a set of simultaneous equations governing the joint states of I_k 's constituents.... In a symmetric nexus, Λ will contain variables referencing the effective properties of each member of the nexus. Λ expresses a set of constraints on the joint states of the constituents of I_k in terms of those variables and constants

Additionally, we hypothesize that the bound members within I_k are encapsulated within interfaces. Their interfaces consist of their own receptivities, through which they holistically receive the constraints in their receptive fields, and their own signals, their effective properties, through which they place elements in the constraint structure on the total state of the nexus. **These interfaces create an information structure within the nexus.**

The above total characterization of a natural individual is fairly substantial from a naturalistic point of view, and it provides some guidance regarding physical world indicators that might be evidence of natural individuality. When searching for natural individuals, this characterization suggests that we should view systems in the physical world as systems of information. For something to be a good candidate for natural individuality, the information system should meet the following conditions:

- 1. Base case: It should be clearly fundamental like a basic particle
- 2. Inductive case:
 - It should be divisible into constituents that are natural individuals themselves, and at least one of which is not known to be in a determinate state, considered independently from the system.

 Its constituent structure should instantiate a system of information based constraints satisfying the conditions on Λ. (Rosenberg 2004, 286; bold added)

Moving from TNI to the Integrated Information Theory (IIT) introduced by Tononi, we see a theory which is almost perfectly structured to be a physical expression of the TNI metaphysics, fitting the description above:

- IIT has panexperiential consequences exactly as predicted in *APFC*, chapter 5.
- IIT makes its predictions based on treating physical systems as information systems as required in *APFC*, chapter 13.
- IIT's core metaphysical notion is mutual causality, mirroring the basic thesis of TNI.
- The elaborated structure of mutual causality is integration of information, which is a perfect example of implementation of the conditions on Λ described above as a set of simultaneous equations describing constraints on the joint states of signal sharing elements.

By joining TNI and IIT, one gets both a metaphysics and a physics for understanding the presence of consciousness in our world: *why* is it present; *where* is it present; and *how much* is present. TNI adds to IIT a deeper but still naturalistic explanation of why integrated information is experiential. It also adds a further physical constraint on where experience exists, requiring integrated information plus at least one component element that would have a physically indeterminate state, considered independently of the integrated system. In turn, IIT adds to TNI a physically instantiable quantity, Φ , whose calculation can potentially provide answers to many detailed questions about the presence and amount of experiencing in specific places in the world.

Baars's Global Workspace Theory (GWT) helps with a further issue. As panexperientialist views, one of the challenges both TNI and IIT share is the need to separate experiencing from cognition. In *APFC* (chapter 5), I draw the obvious implication: we should distinguish between experiencing and consciousness, where the former would represent a kind of raw experiential protoconsciousness existing broadly in nature and where the latter names a species of experiencing shaped specifically by cognition to have certain special attributes we would associate with consciousness (such as representational meaning and awareness).

The Global Workspace theory conjoined with IIT describes the specific cognitive features that shape experiencing into consciousness: a global work-space of integrated information made consistent between long-term memory
and momentary perception, then broadcast for coordinated and global control of the organism. This functionality adds intentionality to experience through preparation of attention and bodily control, it makes room for concepts, and it describes conditions for an individual whose subjective moments are thicker in time than the physical moments of its components. It also helpfully describes conditions for consciousness that could be instantiated by nonanimal creatures or creations as well as by humans and other animals.

The Dynamic Core/Thalamo-Cortical Theory then adds to GWT a plausible biological account of how a global workspace could be implemented in a mammalian brain. It describes extensive two-directional links between cortical areas and the thalamus, well-designed for integrating and broadcasting information of the type described by Baars's GW.

6.7 Understanding Deeply: From 'Big Mysteries' to 'Puzzle Solving'

We should recognize that a paradigm—a set of tools for thinking broadly about questions of consciousness and causality—is emerging here that puts a synoptic understanding of the big questions within our grasp. I believe we are on the verge of being able to shift discussions from 'Big Mysteries' like, "Why does consciousness exist?" and "How does consciousness relate to the physical?" to interesting puzzles in which we can apply tools from our paradigms to get interesting answers to specific questions.

For example, in the summer of 2012, at a philosophical workshop on the combination problem organized by David Chalmers, Keith Turausky proposed a powerful way to think of the intrinsic properties at the basis of the carrier pyramid. He suggested the fundamental qualities may be simple only in the way that, say, white noise or white light is simple. White noise is a relatively simple quality that masks a complex structure, which superposes all possible sounds. It is like playing twenty thousand tones all at the same time, which manifests as one single wave form perceived as qualitatively simple.

Turausky's idea is that the base quality of the Russellian universe is like white noise or white light: a simple quality containing within it the superposition of all possible qualities. Imagine at the basis of the universe's quality space there is a 'fundamental tone,' similar in its relation to other qualities as white noise is to other sounds, in that they can be extracted from the universal waveform given the right filters. For TNI, we should imagine the filtering function of the receptive connections to be like a refracting lens, able to extract or reform individual qualities from the 'fundamental tone' similarly to how a prism refracts white light into component colors.

Mathematically, the analogy is to wave transformation. Every time a set of extracted qualities from lower-level individuals pass through a receptive connection at a higher level, it is as if the intrinsic qualities behave as though they were waves, and a Fourier transform on the shape of the input waves occurs, resulting in a wave of a new shape. I am going to use this comparison to wave forms a lot and want to be clear about what is being said and why. I am not saying experiential qualities *are* waves. I am saying the logic of how waves behave and combine is a useful model for the logic of how experiential qualities behave and combine: they are similar in essential behaviors. But a representation is not the thing represented, and just because we may borrow the logic of wave forms and find it useful or accurate, does not mean the thing *is* a wave.

Combinations of individuals into higher-level individuals require combinations of qualities into new qualities, in a way similar to how combinations of waves create waves with new shapes of simpler or more complex kinds. For example, a set of complicated waves may combine to make a simple straight line *which is a new thing*, and the original waves are not 'components' of it in a compositional way. Similarly, qualities may combine to make a new quality, and not be components of it in a compositional way.

If we assume each quality superposed in Turausky's 'fundamental tone' is high dimensional, we arrive at a picture where different qualities (represented as waves) can end up with peaks and crests in some dimensions offsetting each other, but not those in others, so when combined the resulting qualities end up being 'flat' in different dimensions from one another. This picture implies a mechanism for getting very simple qualities from multiple complex qualities, and producing qualities that are very different from one another because they express different dimensions of the fundamental tone and in different magnitudes. The combination problem is then simply a puzzle and not a mystery: How many dimensions does it take to describe the fundamental tone? What magnitudes are appropriate for describing the peaks and crests in each dimension? What are the transform rules and the matrix addition rules? What qualities would be experienced given different expressions of the tone? If we adopt the Russellian-type framework described in APFC, it's not unreasonable to assume physics is studying this problem 'from the outside' and will eventually produce the core of an answer in the form of a single high-dimensional fundamental physical property from which other properties can be derived in the way described.

This provides guardrails on the kind of experiential transmutation law panpsychists need to describe the evolution of experience from the fundamental to the sophisticated. To show this, I can retell the basic story from *APFC* with some modifications which incorporate the insights above.

- 1. At the bottom—level zero—there is a Turausky Fundamental Tone, the white noise of the universe, in which "twenty thousand experiential tones" are superposed into a single experiential simple. This contains within it the potentials for all possible carriers of effective constraint (physical properties when viewed from the outside). Within the framework of TNI, we view Turausky's Fundamental Tone as the carrier of a universal level-zero effective property, containing within it the potential for all the level-one effective properties in the world. It is the canvas of causation on which the picture of the colorful world is painted (see *APFC*, 143), the abstract metaphysical background against which the jewel of the actual world is set (see *APFC*, 152).
- 2. At the bottom—level zero—there are also experiential quantums. These are level-zero receptivities. In a modification to the framework in *APFC*, let us now say that the carrying capacity of each level-zero receptivity is one. Each can bind to the Fundamental Tone to extract one carrier of a level one effective property—and no more—which will be in an indeterminate state.
- 3. The binding of level-zero receptivities to the Fundamental Tone is therefore the creation of a level one individual—that is, a fundamental effective property (e.g., spin, color)—in an indeterminate state, through extraction of its carrier from the Fundamental Tone. The analogy should be the extraction of a component wave form a compound wave. Because this effective carrier is bound into a receptive nexus alone, without other effective properties, it is unconstrained, and so all its potential values exist in that nexus (see *APFC*, 189). This is why the property is said to be in an indeterminate state.
- 4. Producing property complexes from these level-one properties—that is, producing particles—requires a level-two receptivity to bind the level-one individuals. This is the recursive rule from *APFC* being applied, but we now have IIT as a constraint. The binding of the three fundamental properties together in a particle is a fundamental act of information integration. Let's say the particle receptivity binds three properties—color, charge, spin—inside its nexus, and there are only these three property instantiations in the world. The experiential quantum of the higher-level receptivity must be three, no more and no less, since each level-one individual had an experiential quantum of one. If this higher-level binding creates determinacy, the causality will produce a unified experience-type with a 3-quantum measure on the experience, existing at this higher, determinate level.
- 5. This presents another puzzle to be solved. Conversations with Patrick Lewtas at the same workshop have convinced me that, since we have here a

fundamental theory, we should now be looking for a conservation principle, an integrative law that would be expressed as a function from component qualities to (potentially) more sophisticated qualities which relates a quantity of experience in an *aggregate* to the quantity of experience in an *integration*. What is the form of this law? This is a *puzzle*, not a *mystery*.

It follows from this description that it is perfectly permissible for new kinds of qualities to emerge from level to level—similar to radically new shapes of waves that come from the combining of component waves.

Pulling back from the details, what is remarkable here is the apparent collective achievement of a synoptic hierarchy of theory which can tie together physics, biology, cognition, introspective experience, and integrated information in a way that shows they are different perspectives on the same thing—the causal nexus itself—and it points to ways we can move back and forth between these perspectives using a few basic first principles. I believe the Synoptic Pyramid can take us further—much, much further—toward an integrated understanding of many associated philosophical and scientific puzzles than competing paradigms.

Note

1. In the past I've called this a *possibility filter*, but I think potentiality filter is more appropriate, given that it acts on the potentialities of nomic individuals.

Bibliography

- Baars, Bernard J. A Cognitive Theory of Consciousness. Cambridge: Cambridge University Press, 1988.
- Baars, Bernard J. "The Conscious Access Hypothesis: Origins and Recent Evidence." *Trends in Cognitive Sciences* 6 (2002): 47–52.
- Edelman, Gerald. M. and Giulio Tononi. A Universe of Consciousness. New York: Basic Books, 2000.

James, William. The Principles of Psychology. New York: Dover, 1890/1950.

Rosenberg, Gregg. A Place for Consciousness. Oxford: Oxford University Press, 2004.

Tononi, Giulio. PHI: A Voyage from the Brain to the Soul. New York: Pantheon Books, 2012.

PART III

PANPSYCHISM AND THE COMBINATION PROBLEM

7

The Combination Problem for Panpsychism

DAVID J. CHALMERS

7.1 Introduction

Panpsychism, the view that fundamental physical entities have conscious experiences, is an exciting and promising view for addressing the mind-body problem. I have argued in "Panpsychism and Panprotopsychism" (chapter 1 of this volume) that it promises to share the advantages of both materialism and dualism and the disadvantages of neither. In particular, it can respect both the epistemological intuitions that motivate dualism and the causal intuitions that motivate physicalism.¹

Nevertheless, panpsychism is subject to a major challenge: the combination problem. This is roughly the question: How do the experiences of fundamental physical entities such as quarks and photons combine to yield the familiar sort of human conscious experience that we know and love?

The most influential formulation of the combination problem was given by William James in *The Principles of Psychology* (James 1890). In criticizing "mind-dust theory," on which mental states are held to be compounds of elemental mental states, James made the following observations:

Where the elemental units are supposed to be feelings, the case is in no wise altered. Take a hundred of them, shuffle them and pack them as close together as you can (whatever that may mean); still each remains the same feeling it always was, shut in its own skin, windowless, ignorant of what the other feelings are and mean. There would be a hundred-and-first feeling there, if, when a group or series of such feelings were set up, a consciousness belonging to the group as such should emerge. And this 101st feeling would be a totally new fact; the 100 original feelings might, by a curious physical law, be a signal for its creation, when they came together; but they would have no substantial identity with it, nor it with them, and one could never deduce the one from the others, or (in any intelligible sense) say that they evolved it.

Take a sentence of a dozen words, and take twelve men and tell to each one word. Then stand the men in a row or jam them in a bunch, and let each think of his word as intently as he will; nowhere will there be a consciousness of the whole sentence. We talk of the "spirit of the age," and the "sentiment of the people," and in various ways we hypostatize "public opinion." But we know this to be symbolic speech, and never dream that the spirit, opinion, sentiment, etc., constitute a consciousness other than, and additional to, that of the several individuals whom the words "age," "people," or "public" denote. The private minds do not agglomerate into a higher compound mind.

James is here arguing that experiences (feelings) do not aggregate into further experiences, and that minds do not aggregate into further minds. If this is right, any version of panpsychism that holds that microexperiences (experiences of microphysical entities) combine to yield macroexperiences (experiences of macroscopic entities such as humans) is in trouble.

In recent years, there has been a small groundswell of activity on panpsychism, and in particular there has been a small groundswell of activity on the combination problem. The problem was given its name by William Seager (1995) and was given an especially sharp formulation by Philip Goff (2009). Proposals for addressing it have been presented by Sam Coleman (2012; 2014; this volume), Goff (2009a; 2011; this volume), Gregg Rosenberg (2004; 2014), Seager (2010; this volume), and others.² It is fair to say that no proposed solution has yet gained much support, however.

This chapter is an attempt at a systematic treatment of the combination problem. I distinguish a number of aspects or versions of the problem. I discuss various ways in which the combination problem can be turned into an argument against panpsychism. I then try to systematically lay out the options for dealing with the combination problem, examining their advantages and disadvantages.

A reasonable goal here is to either solve the combination problem or prove that it cannot be solved. I cannot say that I have achieved either of these objectives in this essay as it stands, but I hope to at least clarify the issues enough to help others make progress.

7.2 Terminology

First, some terminology. Most of this terminology is drawn from "Panpsychism and Panprotopsychism" (chapter 1 of this volume), so the presentation here is much briefer than the presentation there. Interested readers may well find it useful to read the other chapter first, though the current chapter is self-contained in principle.

Microphysical properties and entities are the fundamental physical properties and entities characterized by a completed physics. Phenomenal properties are properties characterizing what it is like to be a conscious subject. Microphenomenal properties are the phenomenal properties of microphysical entities. Macrophenomenal properties are the phenomenal properties of other entities, such as humans. Microphenomenal and macrophenomenal truths are truths about the instantiation of these properties.

Constitutive panpsychism is the thesis that macrophenomenal truths are (wholly or partially) grounded in microphenomenal truths. Nonconstitutive panpsychism is the thesis that macrophenomenal truths are not grounded in microphenomenal truths. The most important form of nonconstitutive panpsychism is emergent panpsychism, on which macrophenomenal properties are strongly emergent from microphenomenal or microphysical properties, perhaps in virtue of fundamental laws connecting microphenomenal to macrophenomenal.

Russellian panpsychism is the thesis that microphenomenal properties are quiddities: the categorical bases of fundamental microphysical dispositions, or the properties that play fundamental microphysical roles. For example, the quiddity associated with mass is the property that plays the mass role (resisting acceleration, attracting other masses, and so on). Numerous philosophers have argued that the nature of quiddities is hidden from us. The Russellian panpsychist holds that quiddities are themselves phenomenal.

Perhaps the most important form of panpsychism is constitutive Russellian panpsychism, on which microphenomenal properties serve as quiddities and also serve to constitute macrophenomenal properties. I argue in "Panpsychism and Panprotopsychism" (Chalmers 2013a) that this view is better-suited than any other form of panpsychism to deal with the problem of mental causation. On this view, microphenomenal properties are causally efficacious in virtue of their playing fundamental microphysical roles, and macrophenomenal properties are causally efficacious in virtue of being grounded in microphenomenal properties. By contrast, nonconstitutive and nonRussellian panpsychism have many of the same problems with mental causation as dualism. Panprotopsychism is the thesis that fundamental physical entities have protophenomenal properties. Protophenomenal properties are special properties that are not themselves phenomenal (there is nothing it is like to have them) but that can collectively constitute phenomenal properties. To rule out standard forms of materialism from counting as panprotopsychism, these special properties must be (i) distinct from the structural/dispositional properties of microphysics and (ii) their constitutive relation to phenomenal properties must reflect an a priori entailment from protophenomenal to phenomenal truths.

Constitutive panprotopsychism is the thesis that macrophenomenal truths are grounded in truths about the protophenomenal properties of microphysical entities. Russellian panprotopsychism is the thesis that protophenomenal properties serve as quiddities. Constitutive Russellian panprotopsychism is perhaps the most important form of panprotopsychism, for the same reasons as in the case of constitutive Russellian panpsychism.

7.3 The Many Combination Problems

The combination problem for panpsychism is: How can microphenomenal properties combine to yield macrophenomenal properties? The combination problem for panprotopsychism is: How can protophenomenal properties combine to yield macrophenomenal properties? I will concentrate especially on the problem for panpsychism, but I will address both.

The combination problem can be broken down into at least three subproblems, reflecting three different aspects of phenomenal states: their subjective character (they are always had by a subject), their qualitative character (they involve distinctive qualities), and their structural character (they have a certain complex structure). These three aspects yield what we might call *the subject combination problem*, the *quality combination problem*, and the *structure combination problem*.

The subject combination problem is roughly: how do microsubjects combine to yield macrosubjects? Here microsubjects are microphysical subjects of experience, and macrosubjects are macroscopic subjects of experience such as ourselves.

An especially pressing aspect of the subject combination problem is the *subject-summing* problem. One can pose this problem by an extension of James's reasoning in the passage quoted earlier. Given 101 subjects, it seems that the existence of the first 100 does not necessitate the existence of the 101st. More generally, given any group of subjects and any further subject, it seems possible in principle for the first group of subjects to exist without the

further subject. If so, then no group of microsubjects necessitates the existence of a macrosubject.

The quality combination problem is roughly: How do microqualities combine to yield macroqualities? Here macroqualities are specific phenomenal qualities such as phenomenal redness (what it is like to see red), phenomenal greenness, and so on. It is natural to suppose that microexperience involves microqualities, which might be primitive analogs of macroqualities. How do these combine?

An especially pressing aspect of the quality combination problem is what we might call the *palette problem*.³ There is a vast array of macroqualities, including many different phenomenal colors, shapes, sounds, smells, and tastes. There is presumably only a limited palette of microqualities. Especially if Russellian panpsychism is true, we can expect only a handful of microqualities, corresponding to the handful of fundamental microphysical properties. How can this limited palette of microqualities combine to yield the vast array of macroqualities?

The structure combination problem is roughly: How does microexperiential structure (and microphysical structure) combine to yield macroexperiential structure? Our macroexperience has a rich structure, involving the complex spatial structure of visual and auditory fields, a division into many different modalities, and so on. How can the structure in microexperience and microstructure yield this rich structure?

An especially pressing aspect of the structure combination problem is the *structural mismatch problem*. Macrophysical structure (in the brain, say) seems entirely different from the macrophenomenal structure we experience. Microexperiences presumably have structure closely corresponding to microphysical structure (this is especially clear on a Russellian view), and we might expect a combination of them to yield something akin to macrophysical structure. How do these combine to yield macrophenomenal structure instead?

There are a few other aspects of the combination problem, corresponding to different aspects of macroexperience that need explaining. There is the *unity problem*: How do microexperiences come together to yield a unified consciousness? There is the *boundary problem* (Rosenberg 1998): How do microexperiences come together to yield a bounded consciousness? There is the *awareness problem*: How do microexperiences come together to yield awareness of qualities? And there is the *grain problem* (Maxwell 1979; Lockwood 1993): How do microexperiences come together to yield homogeneous macroexperiences, such as a homogeneous experience of red, instead of an enormous jagged array of distinct qualities? Some of these problems might be assimilated to earlier problems (the first three plausibly involve aspects of subjective character, the last involves an aspect of qualitative character, and all involve aspects of structure), but it is useful to have them on the table explicitly. It is common for a proposed solution to the combination problem to address only one of these problems: most often the subject combination problem and occasionally the quality combination problem. It should be stressed that a satisfactory solution to the combination problem must address all of these problems. This raises the bar for a solution, as it is far from clear that any single proposal can solve all the problems at once. One might appeal to separate proposals for solving the problems one at a time, but then it is far from clear that these proposals will be compatible with each other. At the very least, any proposed solution to the combination problem should indicate which problems it is addressing, and which problems it is not.

The formulation of the problems above is misleading in one respect. I have typically said, "How do microexperiences come together to yield X," or perhaps "How do microsubjects" or "How do microqualities." However, constitutive panpsychism is not committed to the claim that macroexperience is *wholly* grounded in microexperience. It could be partly grounded in causal or structural relations among the microexperiences, or in other microphysical properties, or even in other quiddities if there are nonphenomenal quiddities as well. We can put all this by saying that constitutive panpsychism requires macroexperiences to be wholly grounded in microexperiences and microphysics, where microphysics is understood broadly to include all of the above. The formulations of the relevant problems can then all take the form "How do microexperiences and microphysics come together to yield X?." With the problems understood this way, the panpsychist has more resources to play with, but the problems still seem very difficult to solve.

There are analogous versions of all the problems for panprotopsychism. We need only replace the appeal to microexperience with an appeal to protoexperience (the instantiation of protophenomenal properties), yielding questions of the form: "How do protoexperiences come together to yield X?" or "How do protoexperiences and microphysics come together to yield X?"

The structure and quality combination problems seem just as hard in this guise. The subject combination problem will take a different form, one that perhaps makes it slightly easier. Although microexperiences presumably have subjects, protoexperiences need not. Panprotopsychism therefore need not appeal to microsubjects, and need not require subjects to combine into other subjects. Still, there remains a substantial challenge in explaining how non-subjects of experience can combine to yield subjects of experience.

More generally, panprotopsychism faces a version of the combination problem that does not arise for panpsychism: How can nonexperiences constitute experiences? Sometimes it is flatly asserted that this is impossible, or it is suggested that it is a general gap between the nonexperiential and the experiential that underlies and explains the gap between the physical and the experiential. I do not think that this is obviously correct: I think one can point to special features of the physical that explain the latter gap (the structural nature of physical truths, for example), and I have not seen any argument for a general nonexperiential-experiential gap that is as powerful as the arguments for an physical-experiential gap. Still, there is at least a significant challenge for the panprotopsychist here.

I conclude that both panpsychism and panprotopsychism suffer from serious combination problems.

7.4 Turning the Combination Problem into an Argument

Can the combination problem be proved unsolvable? That requires, in effect, turning the challenges posed by the combination problem into a conclusive argument against pan(proto)psychism or at least against constitutive Russellian pan(proto)psychism. Of course, conclusive arguments are hard to come by in philosophy, but we can at least examine the arguments that are available.

7.4.1 The Anti-Aggregation Argument

One of James's central arguments against panpsychism in *The Principles of Psychology* appeals to the general thesis that aggregates do not really exist: a view sometimes called nihilism about composition. More precisely, James holds that aggregates do not have objective existence, but exist only for observers who perceive them as such. He writes:

In other words, no possible number of entities (call them as you like, whether forces, material particles, or mental elements) can sum themselves together. Each remains, in the sum, what it always was; and the sum itself exists only for a bystander who happens to overlook the units and to apprehend the sum as such; or else it exists in the shape of some other effect on an entity external to the sum itself. Let it not be objected that H_2 and O combine of themselves into 'water,' and thenceforward exhibit new properties. They do not. The 'water' is just the old atoms in the new position, H-O-H; the 'new properties' are just their combined effects, when in this position, upon external media, such as our sense-organs and the various reagents on which water may exert its properties and be known. (James 1890, 158–59) We might try turning this into an argument as follows:

- (1) If constitutive panpsychism is true, human consciousness is an aggregate.
- (2) Aggregates do not objectively exist.
- (3) Human consciousness objectively exists.
- (4) Constitutive panpsychism is false.

The key premise is premise (2). Its support is the claim that aggregates exist only for observers or only in virtue of their effects. James does not give much support for these claims, however, and they are easy to reject. A more orthodox view holds that aggregate entities such as molecules exist independently of observers and independently of their effects. Of course James's nihilism about composite objects is not indefensible. Still, nihilist theses of this sort are so widely rejected that they do not have much dialectical force in an argument against panpsychism.⁴

There is perhaps some intuitive force to the idea that consciousness has a higher and purer degree of existence than tables and molecules. A related argument (consistent with the framework of my "Ontological Anti-Realism," see Chalmers 2009) proceeds from the claim that conscious subjects exist *determinately* whereas aggregates do not. This argument does not require nihilism and arguably applies more plausibly to conscious subjects than to rocks. Still, the premise that aggregates do not determinately exist is highly controversial, so the dialectical force of the argument remains limited.

7.4.2 The Subject-Summing Argument

The subject-summing argument is suggested by James's argument, quoted at the start of the paper, against combination of feelings and minds. We can formalize an argument roughly as follows. As with all the arguments I present in this section and the next, this formalization largely follows the way that closely related arguments are presented in Goff (2009).

- If constitutive panpsychism is true, the existence of a number of microsubjects with certain experiences necessitates the existence of a distinct macrosubject.
- (2) It is never the case that the existence of a number of subjects with certain experiences necessitates the existence of a distinct subject.

⁽³⁾ Constitutive panpsychism is false.

Strictly speaking, premises (1) and (2) should allow arbitrary microphysical truths to be conjoined with the truths about the subjects, but the simple version conveys the main point. Premise (2) is the key premise. An intuitive case for either version of it can be made along the lines of the quote from James at the start of the chapter. One can also support it using a further conceivability argument.

- (1) For any group of subjects (with certain experiences), it is conceivable that those subjects exist (with their experiences) and no other subjects exist.
- (2) For any group of subjects, if it is conceivable that those subjects exist (with their experiences) and no other subjects exist, then this is possible.
- (3) For any group of subjects (with certain experiences), it is possible that the subjects in *S* exist (with their experiences) and no other subjects exist.

Premise (1) has a reasonable degree of intuitive support. Even when adjusted to allow arbitrary microphysical truths to be conjoined with the existence of the subjects in *S*, it retains considerable support. Premise (2) is an instance of a general conceivability/possibility claim. Of course conceivability/possibility claims can be rejected, but not without incurring substantial costs, and panpsychists who have rejected physicalism in part on the basis of conceivability arguments are not in a good position to do so. So this argument poses a significant challenge to the constitutive panpsychist. I will examine options for answering it later.

7.4.3 The Conceivability Argument

The preceding considerations suggest a more general conceivability argument against constitutive panpsychism, inspired by the conceivability argument against physicalism. Here *PP* is a conjunction of all microphysical and microphenomenal truths about the universe, while Q is a macrophenomenal truth, such as "Some macroscopic entity is conscious".

- (1) $PP\&\sim Q$ is conceivable.
- (2) If *PP*&~*Q* is conceivable, it is metaphysically possible.
- (3) If *PP*&~*Q* is metaphysically possible, constitutive panpsychism is false.
- (4) Constitutive panpsychism is false.

Here premises (2) and (3) are parallel to corresponding premises in the familiar conceivability argument against physicalism (see e.g., Chalmers 2009b).

The distinctive premise is (1). This premise in effect asserts the conceivability of a *panpsychist zombie world*: a world in which microphysics and microexperience is just as it is in our world, but in which no macroscopic entity is conscious. Such a world is populated by *panpsychist zombies*, which are microphysical and microphenomenal duplicates of us without consciousness.

Why believe premise (1)? One might think it has a certain intuitive force, just as does the corresponding premise about the conceivability of microphysicalduplicate zombies. However, one can also support it by appealing to the first premise of the conceivability argument in the last section. If we appeal to the modified version of that premise, saying that for any group of conscious subjects and any microphysical truths, it is conceivable that the microphysical truths obtain and the subjects in that group exist without any other subjects, then premise (1) follows.

One might also support premise (1) in other ways. One could use considerations about the quality combination problem to support it, for example arguing that one can conceive of arbitrary microqualities without distinct macroqualities. One could also use considerations about the structure combination problem to support it, arguing that one can conceive of microphenomenal and microphysical structure without distinct macrophenomenal structure. Many of these principles will also generate direct arguments against panpsychism in their own right, but it is useful to have the argument above in the arsenal.

There is also a conceivability argument against panprotopsychism, which replaces *PP* in the argument above by *PPP*, the conjunction of protophenomenal and microphysical truths. The key premise (1) will now say that *PPP&~Q* is conceivable. Why believe this? One might again think it has intuitive support, though this is far from clear given that we have so weak a conception of what protophenomenal properties are like. Alternatively, it might gain support from a thesis holding that for any nonexperiential truth *N* and any experiential truth *E*, *N&~E* is conceivable. Once again, such a thesis might generate a direct argument against panprotopsychism in its own right, but the argument form above helps clarify the territory.

7.4.4 The Knowledge Argument

Having considered the conceivability argument, it is natural to consider a knowledge argument. We can suppose that inside her black-and-white room, Mary is told all the microphysical facts, and also learns all the microphenomenal facts: She learns what it is like to be a quark, a photon, and so on. Perhaps this is accomplished by giving her versions of those experiences, or by somehow enabling her to imagine them. One might think that in this situation, Mary would still be unable to know what it is like to see red, even given

arbitrary a priori reasoning. If so, one could mount an argument as follows. Here *PP* is as before and *Q* is a truth about what it is like to see red, and "deducible" means "inferrable by a priori reasoning alone."

- (1) Q is not deducible from PP.
- (2) If Q is not deducible from PP, Q is not necessitated by PP.
- (3) If *Q* is not necessitated by *PP*, constitutive panpsychism is false.
- (4) Constitutive panpsychism is false.

There is also a corresponding argument against panprotopsychism that replaces microphenomenal facts by protophenomenal facts, and replaces *PP* by *PPP* above. One might likewise think it intuitive that knowledge of all the protophenomenal facts would not help Mary know what it is like to see red.

I think that these arguments are highly inconclusive, largely because we know so little about what microphenomenal or protophenomenal properties are like. Perhaps once we grasped them, we would understand their connection to experiences of red and to other experiences. Certainly there does not seem to be a general case for premise (1) here that is nearly as strong as the case for the premise involving microphysical truths alone. Still, perhaps such a case might be mounted.

7.4.5 The Palette Argument

I turn next to an argument associated with the quality combination problem, inspired by the palette problem discussed earlier.

- (1) If constitutive panpsychism is correct, macrophenomenal qualities are constituted by microphenomenal qualities.
- (2) If Russellian panpsychism is correct, there are only a few microphenomenal qualities.
- (3) Macrophenomenal qualities are too diverse to be constituted by a few microphenomenal qualities.
- (4) Constitutive Russellian panpsychism is incorrect.

Where the previous arguments were arguments against constitutive panpsychism in both Russellian and non-Russellian varieties, this one is an argument against only the former. Russellian panpsychism requires that microphenomenal properties are all directly associated with a fundamental physical property, and there appear to be only a few of these. Non-Russellian panpsychism, by contrast, can escape the argument by allowing that there is a diverse array of microphenomenal qualities.

The case for the key premise (3) is intuitive and inconclusive as it stands: perhaps we might find a small set of deep underlying qualities with sufficient generality to generate all phenomenal qualities, just as we have done for physical qualities. But this is at least an argument that panpsychists need to address.

7.4.6 The Revelation Argument

The revelation argument is also loosely associated with the quality combination problem and is especially closely associated with the grain problem discussed earlier. Versions of this argument are discussed by Lockwood (1993) and Goff (2006).

- (1) The nature of consciousness is revealed to us in introspection.
- (2) If constitutive panpsychism is correct, consciousness is constituted by a vast array of microexperiences.
- (3) Whatever constitutes consciousness is part of its nature.
- (4) A vast array of microexperiences is not revealed to us in introspection.
- (5) Constitutive panpsychism is incorrect.

Premise (1) is not compulsory, and most materialists will deny it. But the premise nevertheless has a certain intuitive plausibility, and some theorists invoke something like it to argue against materialists. For panpsychists who argue in this way, it is an uncomfortable premise to deny. Premises (2) and (4) are also hard to deny.

Perhaps the best way to respond to this argument is to deny premise (3). One can distinguish the nature of a phenomenal property from the grounds (or realizers or constituters) of an instance of that property. It is a familiar point that a single property can be multiply realized by different grounds in different instances, and it is not clear why the same should not also apply to phenomenal properties. It is then coherent to hold that the nature of a phenomenal property is revealed by introspection although the grounds of a specific instance are not.

7.4.7 The Structural Mismatch Argument

This argument is inspired by the structural mismatch problem discussed earlier: macrophenomenal structure (of consciousness) seems quite different from macrophysical structure (of the brain, say) where constitutive Russellian panpsychism would seem to require that the structures be the same. It is also closely related to the grain problem, which is used (for example by Maxwell and Stoljar) to raise a version of the structural mismatch problem.

We can understand microphysical structure and macrophysical structure as the quasi-mathematical structure of microphysical and macrophysical entities as characterized by physics. Macrophenomenal structure is the structure we find within our phenomenology. In both cases, structure includes both internal structure (the internal geometrical structure of a complex physical entity, the internal structure of a visual field) and what we might call external structure: the structure of spaces within which properties are embedded (the scalar structure of mass, the three-dimensional structure of color space).

The structural mismatch argument can be put in the form of an apparently inconsistent tetrad:

- (1) Microphenomenal structure is isomorphic to microphysical structure.
- (2) Microphenomenal structure constitutes macrophenomenal structure.
- (3) Microphysical structure constitutes macrophysical structure.
- (4) Macrophenomenal structure is not isomorphic to macrophysical structure.

Here (1) is an apparent commitment of Russellian panpsychism, (2) is an apparent commitment of constitutive panpsychism, and (3) is a widely accepted view of the physical. (4) reflects the plausible datum of mismatch between the structure of consciousness and the structure of the brain. When combined with the additional premise saying that (1)-(4) are inconsistent, it follows that constitutive Russellian panpsychism is false.

A corresponding argument against panprotopsychism replaces "microphenomenal" by "protophenomenal" in premises (1) and (2). These premises are then apparent commitments of Russellian and constitutive panprotopsychism respectively, so that the inconsistency of the premises yields an argument against constitutive Russellian panprotopsychism.

Although the structural mismatch argument has received relatively little attention to date, I think it is one of the more powerful arguments against constitutive Russellian versions of panpsychism and panprotopsychism. There are various ways to respond to the argument but doing so is not at all trivial. I consider the argument at some length later in this chapter.

7.5 Noncombinatorial Responses

The most obvious sort of panpsychist response to the combination problem is a *combinatorial* response: show how microexperiences can constitutively combine to yield macroexperiences. But there are also *noncombinatorial* responses, which deny that microexperiences constitutively combine to yield macroexperiences.

The most obvious sort of noncombinatorial response is *emergent panpsy-chism*, which holds that macroexperiences are strongly emergent from micro-experiences and are not constituted by them. This view rejects constitutive panpsychism, so it does not need to give an account of mental combination.

Another noncombinatorial response is *identity panpsychism*, on which macroexperiences are identical to microexperiences. On this view, macroexperiences are already present at the fundamental level and no combination is required. Given that microexperiences constitute themselves, this view is nevertheless a form of constitutive panpsychism.

A third noncombinatorial response is *autonomous panpsychism*, which holds that macroexperiences are autonomous from microexperiences, in that they are neither constituted by, emergent from, nor identical to microexperiences. On one version of this view, microexperiences are emergent from or constituted by macroexperiences. On another version, microexperiences and macroexperiences are both autonomous, with neither depending on the other.

These three noncombinatorial responses contrast with the more familiar *combinatorial panpsychism*, on which microexperiences collectively constitute macroexperiences. It is worth noting that all forms of panprotopsychism are combinatorial. By definition, protophenomenal properties are distinct from but can collectively constitute phenomenal properties.

This taxonomy divides panpsychist responses to the combination problem into four classes: emergent panpsychism, autonomous panpsychism, identity panpsychism, and combinatorial panpsychism. The first three are noncombinatorial responses, while the fourth is a combinatorial response. The last two are forms of constitutive panpsychism, while the first two are forms of nonconstitutive panpsychism.

Each of these four broad classes subsumes various specific sorts of response in turn. In this section I discuss the noncombinatorial responses: emergent panpsychism, autonomous panpsychism, and identity panpsychism.

7.5.1 Emergent Panpsychism

Emergent panpsychism holds that macroexperiences are not grounded in microexperiences but instead are strongly emergent from microexperiences, from microphysics, or from both. Strong emergence involves the emergence of ontologically novel entities that are not grounded in the base entities. On a common conception of strong emergence, the base entities do not metaphysically necessitate the emergent entities, but instead they are connected by contingent laws of nature. On this conception of emergent panpsychism, there will be contingent laws of nature connecting microexperience (or microphysics) to macroexperience.

Emergent panpsychism has the great advantage of avoiding the combination problem. Strongly emergent entities and properties are best construed as fundamental entities and properties, not grounded in the base entities or in other entities. As such, no combination is required (except, perhaps, insofar as we construe the laws connecting microexperience with macroexperience as laws of combination). On this view, macrosubjects are fundamental entities, just as they are according to substance dualism. This allows emergent panpsychism to avoid the combination problem just as substance dualism does.

At the same time, emergent panpsychism shares many of the disadvantages of substance dualism. It suffers from problems of economy, postulating many more fundamental entities in the world. And perhaps more important, it suffers from the problems of mental causation. Because macroexperience is not grounded in microphysics or microexperience, it cannot inherit the causal relevance of either. Given that microphysics is causally closed, it is hard to see how macroexperience can have any causal effects on it. Like substance dualism, emergent panpsychism seems to face an unattractive choice between epiphenomenalism, interactionism, and underdetermination.

Of course this does not mean that emergent panpsychism is not true. It may be that it has other advantages over substance dualism, for example with respect to continuity and elegance. It certainly has the advantage of avoiding the combination problem! But for those (like me) who are interested in panpsychism in large part because it promises to avoid the problems of mental causation, emergent panpsychism seems to sacrifice this motivation.

Many solutions to the combination problem that have been put forward turn out on close examination to be forms of emergent panpsychism. For example, Rosenberg (2004) invokes ontologically primitive 'high-level individuals' that emerge from lower-level individuals. Liane Gabora (2002) invokes fundamental principles for 'amplifying phenomenal information,' in virtue of which macroexperience strongly emerges from microexperience. Giulio Tononi's integrated information theory (Tononi 2008), which puts forward a principle connecting degrees of integrated information with states of consciousness, can also be construed as a form of emergent panpsychism. If we see Tononi's principle as a fundamental law of nature, then it appears that macroexperiences are strongly emergent from certain physical configurations.

Many of these theorists do not deal directly with the problem of mental causation. Rosenberg is an exception: he deals with mental causation by allowing high-level individuals to exert a small amount of downward causation

through interaction with the underlying entities. I think that once it becomes clear that these solutions are subject to the same worries about mental causation as substance dualism, they lose some of their initial attractions. Again, this is not to say that these theories are false. But it does give us motivation to look elsewhere.⁵

7.5.2 Autonomous Panpsychism

Like emergent panpsychism, autonomous panpsychism denies that macroexperiences are grounded in microexperiences. Unlike emergent panpsychism, it denies that macroexperiences are even strongly emergent from microexperiences (or from microexperiences and microphysics). In effect, emergent panpsychism retains a sort of dependence of macroexperience on microexperience, if a dependence weaker than grounding or constitution (an asymmetrical nomological dependence, perhaps). Autonomous panpsychism denies even this weak sort of dependence.⁶ As a result, it is not easy to square the view with a contemporary world view on which everything depends at least weakly on what is going on in physics, but the view is worthy of some attention.

One version of autonomous panpsychism says that microexperiences are grounded in (or constituted by) macroexperiences, so that macroexperiences are fundamental and microexperiences are derivative. On one version of this view, human-level experiences are fundamental, as on certain forms of idealism. On another version, universe-level experiences (experience of the whole universe as a subject) are fundamental: a sort of cosmopsychism. These views have to deal with a reverse version of the combination problem, which we might call the decomposition problem. How does macroexperience give rise to microexperience? For example, how does a single subject give rise to multiple dependent subjects? How do macroqualities yield microqualities, and how does macroexperiential structure yield microexperiential structure? These problems seem just as hard as the original combination problem.

Another version of autonomous panpsychism says that microexperiences are strongly emergent from macroexperiences, while a third version says that neither microexperiences nor macroexperiences depend on the other. On these views (as on emergent panpsychism), both microexperiences and macroexperiences are metaphysically fundamental. As with emergent panpsychism and substance dualism, these views avoid the combination and decomposition problems (at least in their hardest forms), but they face the problem of mental causation. On these views microexperiences and macroexperiences are both fundamental, so it appears that they will compete for causal relevance. Given the causal closure of the microphysical, it appears that we have a familiar choice between epiphenomenalism and overdetermination at the macroexperiential level. Again, this does not show that autonomous panpsychism is false, but it gives some motivation for looking at alternative solutions.

7.5.3 Identity Panpsychism

Identity panpsychism holds that macroexperiences are identical to microexperiences: experiences had by fundamental physical entities. This view requires that macrosubjects are themselves microsubjects, or fundamental physical entities. The view may sound unpromising at first, but versions of it are worth exploring.

The version of identity panpsychism that first comes to mind is what we might call the 'dominant monad' view, by analogy to Leibniz's view on which subjects are identical to a single localized monad. On this view, the subject of our experiences is a single localized fundamental entity: perhaps a single quark somewhere in our brain. The microexperiences of this quark are precisely our macroexperiences. There are obvious worries here about this quark's stability (what happens when it disappears?) and about its causal role (how could its properties play the rich causal role that macroexperiences seem to play?).

Even harder problems arise when the view is combined with Russellian panpsychism, on which microphenomenal properties correspond directly to microphysical properties. For a start, the quark is presumably microphysically like other quarks, so it will also be microphenomenally like those quarks, yielding a vast manifold of subjects of experience, just like me, throughout the brain and throughout the universe. And given the simplicity of the microphysical structure of a quark, it is hard to see how the corresponding microphenomenology could have anything like the complexity of our macroexperience. So unless this view is combined with serious revisions to physics, it is probably best put aside.

Other versions of identity panpsychism are *holistic* in that they invoke fundamental physical entities that are not atomic or localized. One such view combines identity panpsychism with the monistic view that the universe itself is the most fundamental physical entity. The result is *identity cosmopsychism*, in which the whole universe is conscious and in which we are identical to it. (Some idealist views in both Eastern and Western traditions appear to say something like this.) Obvious worries for this view are that it seems to entail that there is only one conscious subject, and that each of us is identical to each other and has the same experiences. There is also a structural mismatch worry: it is hard to see how the universe's experiences (especially given a Russellian view on which these correspond to the universe's physical properties) should have anything like the localized idiosyncratic structure of my experiences. Perhaps there are sophisticated versions of this view on which a single universal consciousness is differentiated into multiple strands of midlevel macroconsciousness, where much of the universal consciousness is somehow hidden from each of us. Still, this seems to move us away from identity cosmopsychism toward an autonomous cosmopsychist view where each of us is a distinct constituent of a universal consciousness. As before, the resulting decomposition problem seems just as hard as the combination problem.

Perhaps the most important version of identity panpsychism is *quantum holism*. This view starts from the insight that in the most common understandings of quantum mechanics, the fundamental entities need not be localized entities such as particles. Multiple particles can get entangled with each other, and when this happens it is the whole entangled system that is treated as fundamental and that has fundamental quantum-mechanical properties (such as wave functions) ascribed to it. A panpsychist might speculate that such an entangled system, perhaps at the level of the brain or one of its subsystems, has microphenomenal properties. On the quantum holism version of identity panpsychism, macrosubjects, such as us, are identical to these fundamental holistic entities, and our macrophenomenal properties are identical to its microphenomenal properties.

This view has more attractions than the earlier views, but there are also worries. Some worries are empirical: it does not seem that there is the sort of stable brain-level entanglement that would be needed for this view to work. Some related worries are theoretical: on some interpretations of quantum mechanics the locus of entanglement is the whole universe (leading us back to cosmopsychism), on others there is no entanglement at all, and on still others there are regular collapses that tend to destroy this sort of entanglement. But perhaps the biggest worry is once again a structural mismatch worry. The structure of the quantum state of brain-level systems is quite different from the structure of our experience. Given a Russellian view on which microphenomenal properties correspond directly to the fundamental microphysical properties of these entangled systems, it is hard to see how they could have the familiar structure of our macroexperience.

The identity panpsychist (of all three sorts) might try to remove some of these worries by rejecting Russellian panpsychism, so that microphenomenal properties are less closely tied to microphysical structure. The cost of this move is that it becomes much less clear how these phenomenal properties can play a causal role. On the face of it, they will be either epiphenomenal, or they will make a difference to physics. The latter view will in effect require a radically revised physics with something akin to our macrophenomenal structure present at the basic level. Then phenomenal properties will in effect be playing the role of quiddities within this revised physics, and the resulting view will be a sort of revisionary Russellian identity panpsychism. The overall moral is that it is difficult for the identity panpsychist to avoid epiphenomenalism on one hand or radical revisions in physics on the other. Still, at least the quantum holist version of the view deserves close examination.

7.6 Combinatorial Responses

The most important class of responses to the combination problems is combinatorial responses, on which microexperiences (or protoexperiences) collectively constitute macroexperiences. Here numerous strategies are available. I will start by considering strategies for dealing with the subject combination problem and will then consider strategies for dealing with the other problems.

7.6.1 Deflate the Subject

Any combinatorial version of panpsychism or panprotopsychism must be at least somewhat deflationary about subjects of experience. If subjects were metaphysically primitive entities, they could not be constituted by more basic entities, and combinatorial views would be ruled out. So these views must deny that subjects are metaphysically primitive entities. Indeed, proponents of these views might argue that the subject-summing argument is generated by a tacit background presupposition that subjects are metaphysically primitive entities. If this is right, then replacing this presupposition with a more adequate view of subjects might hold the key to solving the subject combination problem.

An extreme form of deflationism about subjects is *eliminativism*: the view that there are no subjects of experience. If this view is correct, then there are no macrosubjects and the subject combination problem does not need to be addressed. Many of the great neutral monists (themselves panprotopsychists), such as Mach, James, and Russell at least flirted with this sort of eliminativism. Sometimes this view came down to denying a metaphysically primitive subject (as when Mach rejects an 'ego' with 'real unity' and James rejects a 'soul'), but sometimes the view seems to take the more radical form of rejecting subjects altogether, as Russell does in *The Analysis of Matter* and James does in his work on radical empiricism.

Wholesale eliminativism about subjects is not easy to stomach, especially for someone who is serious about phenomenal properties. These properties are defined as those characterizing what it is like to be a subject. And however they are defined, as properties they presumably need bearers, which might then be taken to be subjects. So wholesale eliminativism about subjects may seem to require eliminativism about phenomenal properties or, at least, a reconception of them as properties of quite different entities.

Furthermore, eliminativism does not really remove all aspects of the subject combination problem. Presumably even an eliminativist will still acknowledge that experiences come in bundles or streams of some sort, so that the experiences previously taken to be mine share a bundle or stream with experiences previously taken to be yours. But now the problem can be reconceived as the bundle combination problem, or the stream combination problem: How can a number of distinct streams add up to a new single stream? It is not obvious that this problem is much easier than the original problem. Perhaps the eliminativist can also deny or deflate the bundling relation, but now the view is taking on even more significant costs.

Less extreme views hold that there are subjects while denying that they are metaphysically primitive: perhaps they are composite entities, or they are derivative entities in some other sense. This view has a number of attractions, and can be independently motivated by puzzle cases involving personal identity over time. Still, this sort of deflationism does not make the subject combination problem go away. We still need an account of how a derivative subject of experience can be constituted by microsubjects, or by microphenomenal/ protophenomenal properties along with microphysics. Such a positive account is not easy to find, but I will consider some options in what follows.

7.6.2 Combinatorial Infusion

An idea that is sometimes mooted is that low-level subjects 'merge' or 'blend' or 'fuse' to yield higher-level subjects. After the merging, the low-level subjects no longer exist in their own right. Only the higher-level subject exists. Seager (2010) calls this 'combinatorial infusion,' on which a combined mental state 'supersedes' the original mental states.

Many questions could be raised about this view, but a basic question is the following: Is the relation between the original subjects and the merged subject a synchronic or a diachronic relation? If it is a synchronic relation, then presumably the low-level and high-level subjects exist at the same time, and we have lost the distinctive aspect of this view whereby the high-level subject supersedes the low-level subject. This version of the view will be faced with the original worries about how a number of subjects could ever synchronically constitute another subject.

Presumably the merging relation is diachronic, then. If so, it is hard to see how it can be a constitutive relation. Diachronic relations are naturally understood to be contingent causal relations, not constitutive relations. Perhaps two subjects at an earlier time can nomologically necessitate the existence of a subject at a later time, but it is hard to see how they can metaphysically necessitate or constitute it. But constitution is what we need for a combinatorial solution to the combination problem.

If we examine the synchronic and constitutive structure of this view, it appears to be a form of noncombinatorial panpsychism. At the later time, there is a macrosubject and macroexperiences that are not constituted by microsubjects and microexperiences that exist at that time. So it appears that this macrosubject is itself fundamental. Either we have a version of emergent panpsychism, perhaps with this subject depending nomologically on underlying physical states, or we have a form of identity panpsychism, where this subject corresponds to a fundamental physical state.

One can bring out the point by asking how the view works as a form of Russellian panpsychism. Here the microsubjects and the microphenomenal properties must correspond directly to fundamental microphysical entities and their microphysical properties. So when a number of microsubjects go out of existence and are replaced by a 'merged' subject, a number of microphysical entities presumably also go out of existence, replaced by a 'merged' entity. This does not happen in classical physics, but it can happen in quantum physics. As Seager notes, when two particles become entangled, there is a sense in which neither exists any longer as a fundamental entity: Instead they have 'merged' into a fundamental entangled entity, of which the original particles are at best aspects.

The Russellian panpsychist could exploit this quantum-mechanical merging for their purposes, but the resulting position is a familiar one. It is a version of the quantum holism discussed under identity panpsychism in the previous section. It has the advantages and disadvantages discussed there (notably the worries about stability of entanglement and about structural mismatch), but it is not really a distinct view. Where constitutive relations are concerned, it is a form of identity panpsychism rather than combinatorial panpsychism.⁷

The challenge for this view is making the case that physics really contains infusion laws that yield infusions of the requisite character and complexity. As before, I think that the quantum holism version of the theory is the version most worth taking seriously, but it is not clear that the problems for that view can be overcome. One might try to find another source of infusion in physics, but I suspect that the worries that apply to quantum holism will probably still apply here.

As before, the merging theorist might reject the constitutive Russellian constraints, so that mental merging need not correspond to physical merging, but only at cost of raising serious worries about mental causation. For example, one could also understand laws of combinatorial infusion as 'bridging' laws governing how multiple microsubjects combine to yield macrosubjects, but then the resulting picture appears to be a form of emergent panpsychism. The overall upshot is that combinatorial infusion is best understood as a version of identity panpsychism or emergent panpsychism (with the associated problems), and not as a version of combinatorial panpsychism.⁸

7.6.3 Phenomenal Bonding

Another suggestion (Goff 2009; this volume) is that microsubjects constitute macrosubjects in virtue of certain phenomenal relations between the microsubjects: phenomenal bonding relations. On this view, the subject-summing argument is generated in part by thinking of microsubjects as being merely related spatiotemporally or causally. Once we acknowledge distinctively phenomenal relations between microsubjects and their phenomenal states, we can see how all this might constitute a macrosubject and macrophenomenal states.

An immediate worry question is how there can be room for a phenomenal bonding relation, at least given a Russellian version of panpsychism. But there is an immediate answer. Microphysics postulates fundamental monadic properties such as mass and charge, but it also postulates fundamental relations such as spatiotemporal relations. Just as with mass and charge, physics seems to characterize the mathematical structure of these relations but not their categorical nature. So just as monadic properties can have monadic quiddities underlying them as their categorical bases, relational properties might have relational quiddities underlying them as their categorical bases. It is then not out of the question that a certain phenomenal relation could serve as the quiddity underlying spatiotemporal relations.

A related idea (along the lines of Gregg Rosenberg's "carrier hypothesis" about causation in *A Place for Consciousness* from 2004) is that causation (or perhaps nomic necessitation) is a fundamental relation that has a phenomenal relation as an underlying quiddity. This version of the view also fits well with a Russellian phenomenal bonding theory. One could also invoke non-Russellian phenomenal bonding theories, but as always these will have trouble accommodating the causal relevance of phenomenal bonding and therefore of macrosubjects and macroexperience.

The biggest question for any phenomenal bonding view is as follows: What is the phenomenal bonding relation? And how could any phenomenal relation holding between distinct subjects (or between phenomenal states of distinct subjects) suffice for the constitution of a wholly new subject?

A natural candidate here is the co-consciousness relation: a relation such that whenever it relates two phenomenal states, they are experienced jointly. When this relation holds among the states of distinct microsubjects, those states will be experienced jointly by a new subject. One question for this view and for other phenomenal bonding views is whether the bonding relation is transitive (as co-consciousness seems to be), so that when one microphenomenal state stands in this relation to two other phenomenal states of two other subjects, all three will be jointly experienced by a single subject. If so, then given the ubiquity of spatiotemporal and causal relations, it looks as if the microphenomenal states throughout the universe may stand in this relation, yielding a single giant subject. If on the other hand, the relation is not transitive and one has distinct subjects for different instances of the relation, then one will have far too many subjects, and it is hard to see how we will get macrosubjects. Perhaps there are intermediate possibilities in which the relation is just nontransitive enough to yield nontrivial macrosubjects, but it is hard to see where this structure will come from.⁹

Perhaps there are intermediate options, but it is not at all easy to see how phenomenal bonding will avoid the Scylla of a universal subject and the Charybdis of fragmentary subjects. To yield human consciousness, we presumably want phenomenal bonding to bond a limited multiplicity of microsubjects associated with the human organism, without bonding these to microsubjects elsewhere. It is not at all easy to see what sort of fundamental microphysical relation has this character. Fundamental spatiotemporal and causal relations do not seem to. Perhaps there are derivative causal relations that have this character (a certain sort of informational integration along the lines of Tononi's hypothesis, perhaps?), but these relations are not themselves fundamental. One might suggest that these derivative relations stand to underlying fundamental relation; but now we have a new combination problem concerning how protobonding relations can combine to yield a bonding relation.

One might also worry about the quality combination problem. The coconsciousness relation does not seem to help much here: Presumably the limited palette of microqualities experienced by microsubjects will also be experienced by macrosubjects, and it is not clear how a rich tapestry of macroqualities will emerge. Perhaps there is another sort of phenomenal bonding relation such that bonded microqualities yield a novel macroquality with a different character, but this relation must go well beyond co-consciousness, and it is not clear how it will work.

It is also far from clear how phenomenal bonding will help with the structure combination problem. Insofar as our underlying phenomenal relation is the categorical basis of spatiotemporal or causal relations, one would expect it to have the same structure as those relations, and one would expect the bonded systems to have structure isomorphic to the corresponding composite spatiotemporal or causal structure. But that is not what we find. So new insights are needed here. Still, I think that phenomenal bonding is one of the more promising approaches to the combination problem. I have not begun to canvas all the potential phenomenal relations available to a bonding theorist above, and it is not clear that there is a decisive objection to all such theories (the structural mismatch objection is perhaps the best candidate). So I think phenomenal bonding theories are well worth attention.

7.6.4 Deflating Awareness

Another approach focuses on the awareness relation that subjects stand in to qualities. This relation plays a particularly crucial role as it is arguable that all conscious experience consists in a subject's awareness of qualities. As such, if we can explain how microexperiences and microphysics constitute each instance of the awareness relation between subjects and qualities, we will have solved the combination problem.

It is easy to think of the awareness relation as a primitive relation, on which case it is hard to see how instances of it could be constituted by more basic entities. So a constitutive panpsychist or panprotopsychist may need to deny that it is a primitive relation and explain how instances of it can be constituted.

As with subjects, an extreme deflationary strategy here involves eliminativism: the denial that there is any awareness in experience. This is the strategy famously taken by James in "Does 'Consciousness' Exist?" (see James 1904). He suggests that in experience we find only qualities, with no subjects and no relation of awareness. This view certainly makes the combination problem easier to solve. It has not proved popular, however. It seems introspectively obvious that we are aware of qualities (indeed, I think we are aware of our awareness of qualities; see Chalmers 2013 for an argument). Further, our awareness of qualities plays a natural role in explaining our knowledge of qualities. We can conceive of a situation with qualities that no one is aware of, but such a situation seems very different from ours.

A more moderate deflationary strategy is to endorse some sort of reductionism about the awareness relation. One sort of strategy is to give a causal or functional analysis of awareness. For example, perhaps to be aware of a quality is to stand in a certain causal relation to instances of it, or perhaps it is to have states that play a certain functional role associated with that quality. Given this much, awareness (as a relation between organisms and qualities, say) might be grounded in physical terms alone, or in terms of physical states plus qualities. A version of this strategy is taken by Coleman (2012), who uses a functional account of awareness along with instances of qualities to ground awareness of those qualities. The obvious objection here is that the same considerations that motivate the rejection of functionalism about experience also motivate the rejection of functionalism about the awareness relation. Awareness involves phenomenology, and there are good reasons to think that no mere functional state can constitute phenomenology. For example, one can conceive of any such functional state in the absence of phenomenology, and in particular in the absence of awareness.

Perhaps there are less deflationary accounts of the awareness relation on which it still can be the result of combination. For example, perhaps awareness in microsubjects could somehow constitute awareness in macrosubjects, or protophenomenal properties involving proto-awareness could somehow constitute awareness. It is far from obvious just how this will work, however.

A view like this has the potential to answer the subject combination problem. Anything that is aware of a quality is a subject, so if this approach can show how brains or organisms stand in the awareness relation to qualities, then it will show how brains or organisms can be subjects. On the other hand, the fact that awareness requires subjects might simply suggest that the awareness combination problem is just as hard as the subject combination problem and is subject to the same sort of worries.

The view does not say much about the quality combination problem: It presupposes qualities rather than explaining them. It has the potential to say something about the structural combination problem, by seeing phenomenal structure as the awareness of complex structured qualities. If awareness of those qualities can be explained, phenomenology will be explained. Still, it is not easy to explain awareness of complex structured qualities starting from a base whose structure is quite different.

7.6.5 Panqualityism

A historically popular form of Russellian monism is what Herbert Feigl (1958) called "panqualityism". This is a view on which the quiddities associated with microphysical properties are *qualities*. Qualities are not phenomenal properties. Rather, they are perceived qualities: the properties we are aware of in experience, such as qualitative redness, greenness, squareness, and so on. Arguably for every quality Q, there is a phenomenal property consisting in awareness of Q, and vice versa.

As such, panqualityism is a form of panprotopsychism. Because qualities are so closely related to phenomenal properties, however, this form of panprotopsychism is closely related to panpsychism. Like other forms of panprotopsychism, it can also be seen as a sort of neutral monism. Indeed, something like this seems to have been the preferred view of neutral monists such as Mach, James, and Russell. It has recently been revived in this guise by Sam Coleman (2014).

I have discussed panqualityism at length in "Panpsychism and Panprotopsychism", so I will discuss it only briefly here. Panqualityism (like other forms of panprotopsychism) has the advantage that it has no microsubjects at the basic level, so it avoids James's subject-summing problem. Still, as the view stands, it seems to leave all three main strands of the combination problem open. It is unclear how microqualities can constitute a macrosubject, or how they can constitute macroqualities, or how they can constitute the structure of macroexperience. One needs one of the other solutions to handle each of these issues.

To handle subjects, the historical neutral monists appealed to deflationism (perhaps eliminativism) about subjects and deflationism (perhaps eliminativism) about awareness. More recently, Coleman has appealed to functionalism about awareness here. I think the objections in the previous sections apply strongly here. For example, one can use a conceivability argument (as I do in "Panpsychism and Panprotopsychism") to make the case for an explanatory gap between qualities and awareness, and so between qualities and experience.

In addition, panqualityism does not have much that is distinctive to say about the quality combination problem or the structure combination problem, though perhaps it could adapt elements of other proposals here (and see Coleman this volume). Overall, it seems to me that while panqualityism is an interesting view, it is not obviously more promising than panpsychism in addressing the combination problem.¹⁰

7.7 The Quality Combination Problem

So far I have focused mainly on the subject combination problem. I turn now to the quality combination problem. How do microqualities combine to yield macroqualities? And what in particular of the palette problem: the worry that a small palette of fundamental microqualities cannot generate the vast array of macroqualities that we find in experience?

Qualities here need not be understood as perceived qualities, as in the previous section. Qualities in that sense need not be instantiated in experience. What is instantiated are phenomenal qualities, which involve awareness of perceived qualities. So what we really need to explain is how a small palette of microphenomenal (or protophenomenal) properties can generate awareness of a vast array of macrophenomenal properties. It is quite plausible that principles for combining perceived qualities will play a role in explaining principles for combining phenomenal qualities, but the matter remains open. So I will think about combination both for perceived qualities and for phenomenal qualities.

The initial issue here is whether qualities can combine to yield other qualities at all. We understand how this can work when perceived qualities are coinstantiated. If the same object simultaneously instantiated a degree of redness and a degree of whiteness (at the same location), it will instantiate pinkness. Something similar goes for coinstantiated phenomenal qualities. If the same entity simultaneously is aware of a degree of redness and aware of a degree of whiteness (at the same location), it is plausibly aware of pinkness (at that location). But in general separately instantiated qualities (the redness and whiteness of distinct objects) do not yield a combined quality, and nor do separately instantiated phenomenal qualities. So we need a model of how combination of qualities can work.

This issue may interact with the issue of whether high-level awareness is constituted by low-level awareness. Perhaps such an account can explain how awareness of two distinct qualities by two distinct entities in a complex system can yield awareness of entirely distinct qualities. But it is not at all clear how this will work, especially if we reject highly deflationary accounts of awareness.

What about the palette problem? The two main classes of solutions here are *small-palette* solutions and *large-palette* solutions. Small-palette solutions argue that all macroqualities can be generated from just a few microqualities, if we find the right underlying microqualities with sufficient flexibility and generality. It is far from obvious that such a class can be found, but it is also not obviously out of the question. Small-palette solutions are very much subject to the previous problem of how quality combination works, however.¹¹

Large-palette solutions suggest instead that the full range of macroqualities are included among the microqualities. So there are microqualities associated with different colors, sounds, smells, tastes, and so on. A sufficiently rich largepalette solution might eliminate the need for quality combination altogether, thereby removing the problem of how quality combination works, or at least reducing it to the issue of how macrosubjects can inherit (awareness of) qualities from microsubjects.

The cost is that the plethora of qualities raises familiar problems of mental causation. On a Russellian view, microqualities are causally efficacious in virtue of serving as quiddities for microphysical properties. Given that there are only a few fundamental microphysical properties and one quiddity for each of these, there can be only a few microphenomenal quiddities. So only a few microqualities can be causally efficacious, and the rest will be epiphenomenal.

A large-palette proponent might suggest that microphysical properties can be multiply realized by many different quiddities, but this greatly complicates the simplicity of the standard Russellian view.¹² The suggestion requires that the apparent simplicity of physics (with a small number of fundamental properties and laws) is in fact concealing a much more complex underlying level with a vast multiplicity of fundamental properties and fundamental laws, all connected in such a way to yield the appearance of simplicity. Alternatively the large-palette proponent might reject the Russellian view and deny that microqualities are quiddities, but then they will need another way to make the microqualities causally efficacious. If they allow the microqualities to interfere with microphysical dynamics, this will tend to lead back to a quasi-Russellian view with a much more complicated dynamics. Large-palette solutions seem once again to be stuck with either a form of epiphenomenalism or radical revisions to the fundamental dynamics of the physical world.

7.8 The Structure Combination Problem

What about the structure combination problem: How can microphenomenal and microphysical structure yield macrophenomenal structure? Recall that the structural mismatch argument was presented earlier as an apparently inconsistent tetrad of claims. With a little elaboration we can turn this tetrad into a direct argument against constitutive Russellian panpsychism.

- (1) If Russellian panpsychism is true, microphenomenal structure is isomorphic to microphysical structure.
- (2) If constitutive panpsychism is true, microphenomenal (and microphysical) structure constitutes macrophenomenal structure.
- (3) Microphysical structure constitutes only macrophysical structure.
- (4) If microphenomenal structure is isomorphic to microphysical structure, then any structure constituted by microphenomenal structure (and microphysical structure) is isomorphic to a structure constituted by microphysical structure.
- (5) Macrophenomenal structure is not isomorphic to macrophysical structure.
- (6) Constitutive Russellian panpsychism is false.

Here structure is understood as quasi-mathematical structure involving both internal complexity of states and the quality spaces that they fall into.

Premise (1) is a consequence of the thesis that the quiddity associated with a microphysical property is isomorphic to that property. For example, if mass has a scalar structure, the associated quiddity (what plays the mass role) has a scalar structure. If charge has a binary structure, the associated quiddity (what plays the charge role) has a binary structure. On many Russellian views, microphysical properties such as mass and charge are identical to the associated phenomenal (or protophenomenal) quiddity, in which case they are guaranteed to have the same structure. But even if the two are distinct, one can still expect that in order to be able to play the mass role, a quiddity must have the scalar structure associated with mass.

Premise (2) is close to being true by definition, and premise (3) is highly plausible (perhaps even also true by definition). Premise (4) is in the face of it a plausible general principle about structure. Premise (5) is also highly plausible: The macrophenomenal structure of my visual field is prima facie very different from the macrophysical structure of my brain, and it will often (for example in cases of illusion) be quite different from the macrophysical structure ture of other parts of the world.

The argument is not irresistible, of course. Premise (1) might be denied by someone who holds that quiddities can have surplus structure over and above that of the associated microphysical properties. For example, where mass has a simple scalar structure, perhaps an associated phenomenal quiddity might involve awareness of a certain degree of redness, which has a more complex relational structure due to the role of awareness. One could also say in the reverse direction that microphysical properties have surplus structure that microphenomenal properties do not, perhaps because some but not all microphysical properties have phenomenal quiddities. Still, given a Russellian view, it is not easy to see how these structures could be so different that they yield the vast differences between macrophysical and macrophenomenal structure.

Premise (2) could be denied by someone who says that macrophenomenal structure is constituted by microphenomenal (and microphysical) *qualities,* where these qualities go beyond microphenomenal structure. For this view to help with the problem, specific microphenomenal qualities (phenomenal greenness, say) will have to make a difference to the resulting macrophenomenal structure, so that the latter does not straightforwardly correspond to microphenomenal structure alone. It is not easy to see how this nonstructural factor at the microlevel could make a structural difference at the macrolevel, however.

Premise (3) is true by definition on one reading, where the macrophysical is understood as whatever is constituted by the microphysical. One could deny the premise by understanding the macrophysical more narrowly, however, as I will discuss shortly.

Premise (4) appears to be a plausible principle about structure. It might be denied by someone who holds that, although microphenomenal and microphysical structure are isomorphic, the rules of composition that apply to the former differ from the rules of composition that apply to the latter. It is not
easy to see how this works, however. If microphenomenal and microphysical properties are identical (because mass is identical to the phenomenal property that plays the mass role), it is especially hard to see how a single set of properties could be subject to distinct modes of composition. Even if they are merely isomorphic (because mass is isomorphic to the phenomenal property that plays the mass role), it is hard to see how the two could compose so differently. I return to this issue shortly.

Finally, one could deny premise (5), holding that macrophenomenal structure mirrors macrophysical structure. One route here holds that we are mischaracterizing macrophenomenal structure. Stoljar (2001) suggests that the apparent structure of the visual field is not part of the structure of an experience but only part of the structure represented by the experience. Still, it is plausible that an experience's representational content is itself part of its structure. Even on a representational view, it is plausible that experiences can be similar or different to each other in a manner isomorphic to the way that their representational contents are similar or different to each other: An experience of red₃₁ is similar to an experience of red₃₂ but dissimilar to an experience of green₃₁, just as red₃₁ is similar to red₃₂ but dissimilar to green₃₁. So the relevant structure seems at least to be an aspect of macrophenomenal structure. If so, premise (5) remains plausible.

Another way to deny premise (5) is to hold that there exist macrophysical structures that are isomorphic to apparent macrophenomenal structures: spatial and qualitative replicas of the visual field, for example. These replicas might exist somewhere in the brain, as a physical-sense-datum theorist or a topographic map theorist might hold, or they might exist in the external world, as a naive realist might hold. Still, the macrophysical structure of topographic maps is sufficiently far from that of the visual field to cause problems for the first view, and cases of illusion and hallucination cause obvious problems for the second view.

Perhaps the best way to respond to the argument is to say that it equivocates on 'macrophysical structure.' We might say that *narrowly macrophysical* structure is macroscopic structure characterized in terms of physics: for example, in terms of space, time, mass, charge, and so on. *Broadly macrophysical* structure is any structure constituted by microphysics: for example, chemical, biological, and computational structure. Then a panpsychist can say that premise (3) is true only of broadly macrophysical structure. It is true by definition that microphysical structure constitutes only broadly macrophysical structure, but it is not true that it constitutes only narrowly macrophysical structure, as it constitutes structures that are broadly but not narrowly microphysical. On the other hand, premise (2) is true only of narrowly macrophysical structure: The structure of consciousness is not isomorphic to the spatiotemporal and other narrowly macrophysical structure of the brain, but it may well be isomorphic to other sorts of macrophysical structure there.

Most obviously, one can suggest macrophenomenal structure is isomorphic to certain *information* structure in the brain. For example, the structure of the visual field corresponds to a structure of visual information represented in the brain. I took a version of this line in *The Conscious Mind* (Chalmers 1996). I think something like this has to be the best option for the panpsychist: It seems clear that the structure of the visual field corresponds to information structure in the brain and not to spatial or qualitative structure. The question is whether this line can be made to work.

It is not easy to see how this line can work for a constitutive Russellian panpsychist. From the perspective of physics, high-level information structures are derivative aspects of a more encompassing and more basic narrowly macrophysical structure. We might expect that on a constitutive Russellian view, macrophenomenal properties would have this more basic structure rather than the somewhat arbitrary informational structure. One can bring this out as follows.

On a Russellian view of physics, it is natural to hold that just as there are microquiddities associated with microphysical properties (such as mass), there are macroquiddities associated with narrowly macrophysical properties (such as macroscopic mass). It is also natural to hold that when certain microphysical properties constitute a macrophysical property, the corresponding microquiddities constitute the corresponding macroquiddity. For example, when microphysical mass (one microgram, say) constitutes macrophysical mass (one gram, say), the microquiddity of the former (one unit of pain, say) constitutes the macroquiddity of the latter (one thousand units of pain, either separately or together?). Because the macroquiddity corresponds so closely to the macrophysical property, we should expect them to have isomorphic structure for reasons discussed under premise (1). On a constitutive Russellian view, it is natural to hold that these macroquiddities are macrophysical properties, which will then be isomorphic to narrowly macrophysical properties.

At this point, the constitutive Russellian panpsychist may say there are both *narrow macroquiddities*, the quiddities of narrowly macrophysical properties, and *broad macroquiddities*, the quiddities of broadly macrophysical properties, with different macrophenomenal properties playing both roles. Then the macrophenomenal properties we experience might be broad macroquiddities: quiddities of informational properties, for example. This view naturally goes with the view that while microphenomenal and narrow macrophysical properties are highly natural and play a special role in physics, the macrophenomenal properties we experience (like broadly macrophysical properties) are less natural and more arbitrary from the point of view of physics.

Still, it remains unclear just why phenomenal microquiddities should give rise to broad phenomenal macroquiddities. It also remains unclear how these broad phenomenal macroquiddities relate to narrow phenomenal macroquiddities. In particular how can these two sorts of macroquiddity stand in the constitutive relation that is plausibly required to avoid causal exclusion worries?

I suggested in *The Conscious Mind* (Chalmers 1996) that the principles of phenomenal composition might more closely reflect the constitution of information than the constitution of standard macrophysical structure. Again, I think that something like this is perhaps the only viable line for a panpsychist or panprotopsychist. But it is not at all clear why or how phenomenal composition could work this way while still being a sort of constitutive composition. Certainly one could articulate laws of informational structure for phenomenology, but it is not easy to see how these will be metaphysically necessary rather than brute nomic principles.

In any case, if we are looking to either solve the combination problem or to prove it unsolvable, I think the structural mismatch problem is a promising place to focus. It may be that reasoning along the lines I have given here can be made more rigorous to exclude all possible solutions; and it may be that tightening up the reasoning will reveal the avenues that a panpsychist or panprotopsychist may exploit. In any case, it is clear that the structural mismatch argument is a significant challenge that all Russellian monists must answer.

7.9 Conclusion

What, then, are the prospects for solving the combination problem? On my view, the avenues that seem to be perhaps the most worth exploring are phenomenal bonding or quantum holism (to solve the subject combination problem), small qualitative palettes (to address the quality combination problem), principles of informational composition (to address the structure combination problem), and a somewhat deflationary account of awareness of qualities to tie all these aspects together. It is not at all clear whether these ideas can work together in such a way that all the combination problems are solved at once, however.

After a close analysis of the many aspects of the combination problem and the limited resources for solving them, it is easy to be pessimistic about the prospects for a solution. What emerges is that panpsychism and panprotopsychism, at least in their constitutive Russellian form, are subject to extraordinary constraints in finding a theory of consciousness. It is hard enough to find a theory of consciousness that works on dualist terms, where we are allowed to take macrosubjects and macrophenomenal properties as primitive and appeal to numerous contingent psychophysical laws. The Russellian monist is constrained to find a theory whereby macroexperience is constituted by a tiny range of underlying primitive properties and without any further contingent fundamental laws. This is a little like trying to juggle seven balls in the air with both hands tied behind one's back.

It may be that the constraints imposed by the combination problem are so strong that the challenge cannot be answered. Or it may just be that trying to satisfy the constraints will point someone toward the correct form for a fundamental theory of consciousness.

Notes

- 1. This chapter is based on my opening presentation to "Panpsychism on the Reef," a workshop on the combination problem held on Lady Elliot Island in July 2012. Thanks to the audience there for useful discussion. The material on the structural mismatch problem benefited from discussion at the Oslo conference on panpsychism in August 2013. I am grateful to many philosophers for their responses to the first draft of this paper, both in conversation and in print. In some cases I have added some new discussion in response, mainly in footnotes. For written comments I am grateful to John Gregg and Tom McClelland.
- 2. Other recent work discussing the combination problem includes Basile 2010; Blamauer 2011; Dainton 2011; Gabora 2002; Goff 2006; Hunt 2011; Montero this volume; Mørch 2014; Roelofs 2014; Shani 2010; Skrbina 2011; and Strawson 2006b.
- 3. Barry Dainton (2011) calls this problem the "derivation problem."
- 4. Itay Shani (2010) gives a thorough discussion of James's argument here.
- 5. Hedda Hassel Mørch (2014) defends emergent panpsychism by holding that (i) emergent causal relations can be intelligible rather than brute; (ii) macroexperiences have metaphysical priority over the microexperiential parts from which they emerge; and (iii) macroexperiences are the intrinsic natures of certain macroscopic physical systems that have metaphysical priority over their microphysical parts. Mental causation is handled by the observation that macroexperiences are more fundamentally efficacious than their microphysical parts. Challenges for this view include understanding how macroscopic entities and properties can be metaphysically prior to the microscopic entities and properties that cause them; and understanding how and whether intrinsically identical microscopic entities and properties will be causally efficacious or not depending on their macroscopic surrounds. One could handle these challenges by taking the view to be a version of the 'combinatorial infusion' view (discussed in this chapter) with fundamental fused entities in the physics, but the various challenges for that view must then be met.
- 6. In forthcoming work, Jennifer McWeeny argues that the seventeenth-century philosopher Margaret Cavendish was a sort of autonomous panpsychist, holding that everything in the universe is conscious and that consciousness at one level does not depend on consciousness at other levels.
- 7. Seager (this volume) suggests that his combinatorial infusion view can avoid various versions of the combination problem by appealing to laws of combinatorial infusion, which are fundamental laws of nature akin to laws of physics. On a constitutive Russellian panpsychist position, it is natural to hold that the only fundamental mental laws will be mental "realizations" of the fundamental laws of physics. Where physical properties are realized by mental quiddities, then laws connecting those properties will be realized by isomorphic laws connecting the corresponding quiddities. On this picture, any laws of

combinatorial mental infusion must be realizers of a corresponding law of infusion in the fundamental physics, and the infused mental entity will realize an infused entity in the fundamental physics (a holistic quantum system, perhaps). This clearly leads to a form of identity panpsychism along the lines in the text, rather than a form of combinatorial panpsychism.

- 8. In conversation, Tom Nagel has suggested a panprotopsychist version of the infusion view, on which protophenomenal properties yield experiencing subjects as follows. If a fundamental physical entity is sufficiently isolated, its protophenomenal character determines an individual subject. If it is in the right kind of complex system, it instead contributes to determining a more complex (merged or infused) subject necessitated by the system as a whole. This view is in some ways reminiscent of Tononi's integrated information theory, whose exclusion postulate says roughly that a system is conscious if it is not part of (and does not contain) a system with a higher degree of integrated information. Both views seem to have the counterintuitive consequence that consciousness is extrinsic. Extrinsically identical physical systems (with the same fundamental physical and protophenomenal properties) might be conscious or non-conscious depending on the surrounding context.
- 9. Dainton (2011) suggests that a nontransitive view of co-consciousness can help with the combination problem by making it coherent that microsubjects and macrosubjects share experiences, but he does not really address how the relation could be structured to yield a nontrivial structure of macrosubjects.
- 10. A sixth idea to address the subject combination problem, proposed by Luke Roelofs, is that of *mereological inheritance*: Composite entities inherit experiences from subjects that are their parts. Roelofs (2014) proposes this in either a 'conditional' version where the subjects must meet certain further conditions (e.g., being appropriately related) or a 'basic' version where any composite inherits the experiences of its parts (perhaps because all fundamental properties are inherited by wholes from parts). The obvious problem for most versions is that inheritance principles of either sort do not seem to be a priori. It seems that one can straightforwardly conceive of the relevant microsubjects without any inheritance by macrosubjects. If so, then for this view to yield a version of constutive panpsychism, the inheritance principles will have to be strong a posteriori necessities. The same worry applies to other elements of Roelofs's interesting and comprehensive framework for dealing with the combination problem.
- 11. Roelofs (2014) outlines a small-palette view of quality combination in terms of operations whereby microqualities are 'confused' and 'refracted' into macroqualities by high-level cognitive processes (see also Coleman this volume for a related view in terms of 'contamination'). Prima facie this proposal leaves the usual explanatory gaps (one can conceive of the low-level qualities and the physical dynamics without any refraction into high-level qualities), so again it is not easy to see how it works as a sort of constitutive panpsychism without an appeal to strong a posteriori necessities.
- 12. Tom McClelland has suggested a multiple-realization version of a large-palette view to me, while Pat Lewtas (*forthcoming*) has suggested a non-Russellian version.

Bibliography

- Basile, Pierfrancesco. "It Must Be True, but How Can it Be? Some Remarks on Panpsychism and Mental Composition." Royal Institute of Philosophy Supplement 85.67 (2010): 93–112.
- Blamauer, Michael. "Taking the Hard Problem of Consciousness Seriously: Dualism, Panpsychism and the Origin of the Combination Problem." *The Mental as Fundamental*. Ed. Michael Blamauer. Frankfurt: Ontos, 2011.99–116.
- Chalmers, David. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press, 1996.
- Chalmers, David. "Consciousness and Its Place in Nature." *The Blackwell Guide to Philosophy* of Mind. Ed. Stephen Stich and Ted Warfield. Oxford: Wiley-Blackwell, 2003. 102–42.

- Chalmers, David. "Ontological Anti-Realism." *Metametaphysics: New Essays in the Foundations* of Ontology. Ed. David Chalmers, David Manley and Ryan Wasserman. New York: Oxford University Press, 2009. 77–129.
- Chalmers, David. "The Two-Dimensional Argument Against Materialism." *The Character of Consciousness.* Ed. David Chalmers. New York: Oxford University Press, 2010. 140–206.
- Chalmers, David. Constructing the World. New York: Oxford University Press, 2012.
- Chalmers, David. "Panpsychism and Panprotopsychism." *The Amherst Lecture in Philosophy* 8 (2013a): 1–35.
- Chalmers, David. 2013b. "The Contents of Consciousness: Reply to Hellie, Peacocke, and Siegel." *Analysis* 73.2 (2013b): 345–68.
- Coleman, Sam. "Mental Chemistry: Combination for Panpsychists." *Dialectica* 66.1 (2012): 137–66.
- Coleman, Sam. "The Real Combination Problem: Panpsychism, Micro-Subjects, and Emergence." *Erkenntnis* 79.1 (2014): 19–44.
- Coleman, Sam. "Panpsychism and Neutral Monism: How to Make up One's Mind." *Panpsychism*. Ed. Godehard Brüntrup and Ludwig Jaskolla. New York: Oxford University Press, this volume.
- Dainton, Barry. "Review of 'Consciousness and Its Place in Nature." Philosophy and Phenomenological Research 83.1 (2011): 238-61.
- Feigl, Herbert. "The 'Mental' and the 'Physical." Minnesota Studies in the Philosophy of Science Vol. 2. Ed. Herbert Feigl, Michael Scriven & Grover Maxwell. Minnesota: University of Minnesota Press, 1958. 370–497.
- Gabora, Liane. "Amplifying Phenomenal Information: Toward a Fundamental Theory of Consciousness. Journal of Consciousness Studies 9.8 (2002): 3–29.
- Goff, Philip. "Experiences Don't Sum." Journal of Consciousness Studies 13.10-11 (2006): 53-61.
- Goff, Philip. "Why Panpsychism Doesn't Help to Explain Consciousness." *Dialectica* 63.3 (2009): 289–311.
- Goff, Philip. "Can the Panpsychist get Round the Combination Problem?" Mind that Abides: Panpsychism in the New Millennium. Ed. David Skrbina. Amsterdam: John Benjamins, 2009a. 129–35.
- Goff, Philip. "There Is No Combination Problem." *The Mental as Fundamental*. Ed. Michael Blamauer. Frankfurt: Ontos, 2011. 131–40.
- Goff, Philip. "The Phenomenal Bonding Solution to the Combination Problem." Panpsychism. Ed. Godehard Brüntrup and Ludwig Jaskolla. New York: Oxford University Press, this volume.
- Hunt, Tam. "Kicking the Psychophysical Laws into Gear: A New Approach to the Combination Problem." *Journal of Consciousness Studies* 18.11–12 (2011): 96–134.
- James, William. The Principles of Psychology. Boston: Henry Holt, 1890.
- James, William. "Does 'Consciousness' Exist?" Journal of Philosophy, Psychology, and Scientific Methods 1.18 (1904): 477–91.
- Lewtas, Patrick. "Building Minds: Solving the Combination Problem." Forthcoming.
- Lockwood, Michael. Mind, Brain, and the Quantum. New York: Oxford University Press, 1989.
- Lockwood, Michael. "The Grain Problem." *Objections to Physicalism*. Ed. Howard Robinson. New York: Oxford University Press, 1993. 271–92.
- Maxwell, Grover. "Rigid Designators and Mind-Brain Identity." *Minnesota Studies in the Philosophy of Science* 9.9 (1979): 365–403.
- Montero, Barbara. "What Combination Problem? *Panpsychism*. Ed. Godehard Brüntrup and Ludwig Jaskolla. New York: Oxford University Press, this volume.
- Mørch, Hedda. "Panpsychism and Causation: A New Argument and a Solution to the Combination Problem." Diss. University of Oslo, 2014.
- Roelofs, Luke. "Combining Minds: A Defence of the Possibility of Experiential Combination." Diss. University of Toronto, 2014.

- Rosenberg, Gregg. "The Boundary Problem for Phenomenal Individuals." *Toward a Science of Consciousness: The Second Tucson Discussions and Debates.* Ed. Stuart Hameroff, Alfred Kaszniak, and Alwyn Scott. Cambridge: MIT Press, 1998. 149–56.
- Rosenberg, Gregg. A Place for Consciousness. New York: Oxford University Press, 2004.
- Rosenberg, Gregg. "Causality and the Combination Problem." Consciousness in the Physical World: Perspectives on Russellian Monism. Ed. Torin Alter and Yujin Nagasawa. New York: Oxford University Press, 2016. 224–45.
- Russell, Bertrand. The Analysis of Mind. London: George Allen & Unwin, 1921.
- Russell, Bertrand. The Analysis of Matter. London: Kegan Paul, Trench, Trubner, 1927.
- Seager, William. "Consciousness, Information, and Panpsychism." Journal of Consciousness Studies 2.3 (1995): 272–88.
- Seager, William. "Panpsychism, Aggregation and Combinatorial Infusion." *Mind and Matter* 8.2 (2010): 167–84.
- Seager, William. "Panpsychist Infusion." *Panpsychism*. Ed. Godehard Brüntrup and Ludwig Jaskolla. New York: Oxford University Press, this volume.
- Shani, Itay. "Mind Stuffed with Red Herrings: Why William James' Critique of the Mind-Stuff Theory Does Not Substantiate a Combination Problem for Panpsychism." *Acta Analytica* 25.4 (2010): 413–34.
- Skrbina, David. "Mind Space: Toward a Solution to the Combination Problem." *The Mental as Fundamental.* Ed. Michael Blamauer. Frankfurt: Ontos, 2011. 117–30.
- Stoljar, Daniel. "Two Conceptions of the Physical." *Philosophy and Phenomenological Research* 62.2 (2001): 253-81.
- Strawson, Galen. "Realistic Monism: Why Physicalism Entails Panpsychism." Journal of Consciousness Studies 13.10–11 (2006a): 3–31.
- Strawson, Galen. "Panpsychism? Reply to Commentators with a Celebration of Descartes." Journal of Consciousness Studies 13.10–11 (2006b): 184–280.
- Tononi, Giulio. "Consciousness as Integrated Information: A Provisional Manifesto." *The Biological Bulletin* 215.3 (2008): 216–42.

8

What Combination Problem?

BARBARA GAIL MONTERO

That I may understand whatever

Binds the world's innermost core together

—Goethe's Faust, line 382

It would seem that one reason to be a panpsychist—that is, to think that conscious experience more or less pervades the universe—is that panpsychism obviates the need to bridge what some see as the yawing gap between the conscious and the nonconscious: no need to derive mind from brute matter, no need to explain how nonconscious particles give rise to conscious beings, no need to squeeze the experience of ecstasy out of the unfeeling, dreary activations of the orbitofrontal and anterior cingulate cortex. According to the panpsychist, mind is part of the basic fabric of the universe and as such, bridge-building engineers are unnecessary. For those who are convinced that we shall never explain consciousness in terms of nonconscious goings on, panpsychism, then, would seem a good way to go.

But although panpsychism certainly precludes the demand for an explanation of the conscious in terms of the nonconscious, might panpsychists need to accomplish a different, though perhaps just as arduous, engineering feat? According to Daniel Stoljar, Philip Goff, and others, panpsychists need bridges as much as physicalists do since it is no easier to account for how consciousness arises from microconscious phenomena than for how it arises out of microphysical, nonconscious phenomena. This is because although panpsychism, in their eyes, requires experiences at the microlevel to combine into human-level experience, "experiences don't sum," as Goff (2006, 53–59) puts it, or as Sam Coleman paraphrases Goff's view, "you can stack them, but you cannot pool them" (Coleman 2012a, 147).

The difficulty of summing experiences is referred to as "the combination problem," and here I suggest that this difficulty, as conceived of by Coleman

and others, is really not so difficult after all—subjects, in fact, can pool—but more importantly, I shall argue that the problem is ill-conceived. Coleman seems to think that it is incumbent on the panpsychist to show how minds can pool at one level—like the way antifreeze might pool under your car if you've sprung a leak—however, I shall argue that panpsychists should understand the combination of microminds not as involving the type of pooling that leads simply to the formation of larger wholes but rather in line with the relations we find in the world between higher and lower levels of organization: the relationship between microminds and macrohuman-level minds is analogous to the relation between atoms and molecules or macromolecules and organelles or even populations and communities. It is along these lines, the panpsychist ought to hold, that the world's innermost core, in Goethe's phraseology, ultimately gives rise to human consciousness.

One might see my argument as an attempt to solve the combination problem, for I do aim to show that there is no barrier to thinking that microminds can form the subvenient base for higher-level macrominds. However, I prefer to think of my task as a dissolving one, for I see the combination problem as iatrogenic: induced by philosophers in their attempts to cure panpsychism rather than following from panpsychism itself.

8.1 The Combination Problem

William James (1890/1950) presented what is now seen as the classic argument for the view that minds do not combine and thus that panpsychism offers no advantage over physicalism when it comes to bridging the explanatory gap between mind and matter (for discussion of Jame's argument see Coleman 2012a, Shani 2010, Chalmers this volume). According to James, individual minds, no matter how close in proximity, never form a group mind. We may, as he puts it, "talk of the 'spirit of the age,' and the 'sentiment of the people,'" but such talk, he thinks, is not to be taken literally; for on his view, "private minds do not agglomerate into a higher compound mind" (James 1890/1950, 160). Proponents of the combination problem take this up and argue that if individual minds do not agglomerate, it is just as difficult to see how microminds could combine to form a human mind as it is to see how microphysical particles could do so. Hence, the engineering costs for the panpsychist are high—perhaps impossibly high.

A thought experiment presented by Ned Block (1980), which was originally aimed at discrediting functional accounts of mentality, is interpreted by Daniel Stoljar (2006) as making a similar point. Block asks us to imagine that incredibly small aliens have, for reasons known only to them, decided to duplicate the exact functional description of our elementary particles by flying around in hordes of particle-sized spaceships. According to Stoljar, even though the aliens at the controls are conscious, we are at a loss to see how the human-like entities that are realized by this odd alien-spaceship arrangement are conscious. Hence, in Stoljar's eyes, "it seems just as hard to see how one experiential truth can entail another as it is to see how a nonexperiential truth can entail an experiential truth" (Stoljar 2006, 120).

Chalmers (this volume) tells us that, actually, there are (at least) three distinct combination problems that challenge panpsychism: the "subject combination problem," which is the problem of how to combine subjects of experience; the "quality combination problem," which is the problem of how microqualities combine to form macroqualities; and the "structural combination problem," which is the problem of how the structure of the microexperiential yields macroexperiental structure, where structure can be thought of, roughly, not as the microexperiential and macroexperiential entities themselves but rather as the set of relations that such entities instantiate (for example, not the electron itself, but rather the property that electron has of repelling or attracting other electrons). And according to Chalmers, it is incumbent upon anyone proposing a solution to the combination problem to indicate which of these variations is being addressed. But is it?

As methodological principle, it would seem at least sometimes advisable to disregard prescriptions for divvying up philosophical terrain. Agreement on the central questions certainly may make philosophy more like what Kuhn spoke of as normal science, where researchers converge on the basic problems and work within a shared conceptual framework. Yet as I see things, progress in philosophy—or if not progress, then at least exciting developments—are made precisely by those who categorize the issues in entirely new terms.¹ Before the Kantian revolution, for example, part of the philosophical problem space was thought to include the question of where to find causal relations in the world. Progress (or, again, at least exciting developments) occurred when Kant, rather than addressing the standard question, reconceptualized the terrain. Nonetheless, I can say that both in questioning Coleman's argument for the impossibility of subjects combining and in attempting to dissolve the combination problem, I assume that what is thought to combine is either subjects or some other aspect of mentality. Not everyone can be a Kant.

8.2 The Purported Preclusion of Perspectival Pooling

According to Coleman, both the James combination problem and the Block/ Stoljar thought experiment assume that the panpsychist's fundamental constituents, what he refers to as "ultimates," are subjects of experience, and subjects, as he sees it, do not combine: "little minds assembled do not pool into a corporate mind," and if the panpsychist starts with subjects as her ultimates, our conviction that subjects do not pool, "would appear an effective *reductio* of panpsychism" (Coleman 2012a, 141). The panpsychist, Coleman thinks, then must find experience but not subjects in the world's innermost core. And this is what Coleman, a self-described "panpsychist fruitcake," aims to do.² But is it impossible for subjects to pool?

Coleman thinks that it is: "our notion of a mind," he tells us, "like our notion of a subject is precisely the notion of a discrete, essentially inviolable sphere of conscious-experiential goings-on. My mind is separate from your mind, is separate from her mind, and so on" (Coleman 2012a, 145). What is at issue here is not simply that subjects are thought of as discrete, for Coleman readily admits that discrete things can combine. Discrete hydrogen atoms combine with discrete oxygen atoms to make water and when he makes lasagna, he says, the discrete ingredients of the ragout combine to form a unity: "The red wine infuses the tomatoes . . . [t]he sauce leaks into the ground beef . . . [t]he onions become garlicky" (Coleman 2012a, 140). I'm impressed by his culinary skills. But if such combination happens, as Coleman points out, "all across the natural world" (Coleman 2012a, 140), what, then, is at issue in the combination problem? What, then, makes subjects inviolable?

Coleman argues that because the combination of discrete things is prevalent in the natural world, "if there's something distinctively problematic about the combination of phenomenal elements, it must derive from the fact of their phenomenality" (Coleman 2012a, 140). However, if one admits that there are some types of things in the natural world that do not combine—the periodic table, after all, contains a limited number of elements—the implication does not hold since phenomenal combination might be impossible for reasons that have nothing to do with phenomenality. (Or at least that could be so if phenomenal elements have properties other than phenomenal properties. It would be those properties that preclude combination.) So it would seem possible that the combination problem, if there is a combination problem, arises not because of how we are conceptualizing the experiences of subjects at the world's innermost core, but rather because of various other limitations on how things come together. Nevertheless, Coleman thinks the combination of minds is problematic because of a feature specific to minds, a feature that minds have yet red wine and tomatoes lack. What, then, is this specific feature?

The feature that precludes pooling, on Coleman's account, seems to have something to do with the nature of perspectives. Minds can't combine, he says, because acts of introspection will disclose information about what occurs only in your own mind: "Intuitively, phenomenal perspectives—minds, subjects include at a time a discrete set of phenomenally conscious elements, to which an introspective act on the part of one such phenomenal perspective has access." This, at least, is his first pass at explaining why subjects (or perspectives, or minds) can't combine. Mental combination cannot occur because mental combination would imply that we could introspect another's mind, yet we have introspective access only to our own minds. Thus, minds don't combine.

If this is correct, panpsychism that posits subjects at the ground level is not only no better than physicalism at bridging the gap between microlevel phenomena and macrolevel conscious phenomena, it is much worse. Physical particles can at least amalgamate; subjects, on Coleman's view, are irrevocably distinct.

8.3 Partial Overlaps

By claiming that subjects or perspectives cannot pool, Coleman does not mean to deny that you and I could be in qualitatively identical states; when subjects pool, it is not merely "as if we are looking attentively at the same dog from more or less the same angle" (Coleman 2012a, 146). Arguably, if *a* and *b* are in qualitatively identical states, they could still count as numerically distinct individuals if they were to occupy different spatiotemporal locations. Rather, what I understand him to mean in claiming that two subjects cannot share one and the very same perspective, is that they cannot have one and the very same point of view.

Sharing a perspective, for him, is also not simply sharing mental content, for, on his view, there could be beliefs, pains, and perceptions that, though mental, are not necessarily perspectival, they are not necessarily connected to a point of view. And while perspectives, on his view, do not pool, mental content may do so to a degree. He mentions two different situations which he thinks suggest the possibility of partial pooling of mental content. One involves telepathy: "To be precise, *if* telepathy is possible then it seems that some mind-to-mind conditioning is feasible" (Coleman 2012a, fn. 15). If telepathy, on his view, were possible, then mental contents could infuse one another a little bit (somewhat less, I imagine, than what happens after his ragout is left in the fridge for a day or two). The second situation he countenances in which mental connecting the brain's two hemispheres has been severed). In such a case, he thinks, it may be that two subjects, each corresponding to a single hemisphere, both have introspective access to an element of experience.³

Though Coleman does not mention it, it would seem that another and far less fantastical way that mental contents can marry is illustrated by the strange and beautiful case of the conjoined twins Krista and Tatiana Hogan. Krista and Tatiana are craniopagus twins, that is, twins conjoined at the head. However, they are highly unusual, and perhaps unique because not only are their skills fused, but their brains are linked by a complex neural connection, what their neurosurgeon Dr. Doug Cochrane dubs the "thalamic bridge" (Dominus 2011), which allows information to pass from one twin to the other. Arguably, this bridge facilitates experiential pooling, for reportedly Krista, with eyes covered, can somehow move into Tatiana's mind and identify what Tatiana is seeing. Their condition, from what I gather, is not very well understood and the scientific experiments performed on them by their neurosurgeon are as of yet unpublished (see Dominus 2011). However, if this story is accurate, it seems to be an actual example of two individuals not only having overlapping neural realizers of their visual experiences, and not only have overlapping mental content, but occasionally having overlapping visual fields, or perspectives, for the experience of moving into the other's mind is the experience of moving away from one's individual perspective and inhabiting another's.

Are Krista and Tatiana are actually two subjects? Given that they typically seem to inhabit their own perspectives and, as the case is described, it is only with a bit of effort that the one is able to move into the perspective of the other, it seems reasonable to count them as two individuals (though perhaps it might be that the case indicates there are situations in which there is no clear answer to the question of whether there are one or two subjects.) In any event, their situation should make us question Coleman's initial contention that phenomenal perspectives are "by their fundamental nature closed off by each other" (Coleman 2012a, 145) as well as provide some insight into how one individual can have introspective access into another individuals' mind.

8.4 The Impossibility of Total Necessary Overlap

Coleman wants to allow for the possibility that telepathy and split brain cases could facilitate partial overlap of mental content, or qualitative overlap. But partial overlap of mental content, as he sees it, does not entail perspectival overlap. Necessary total overlap of mental content, on his view, does entail perspectival overlap, for his view is that if a and b's mental contents were to necessarily overlap entirely, then a and b are not two different subjects. In other words, according to Coleman, if whenever a reflects on her own internal mental states she necessarily thereby has direct access to b's mental states, then never the twain shall part. With telepathy, he tells us, "despite having some qualitative overlap, our phenomenal perspectives are irrevocably separate" (Coleman 2012a, 146), and with the split brain case, he says that since there would not be total overlap of mental content, we would still have "two discrete minds on our hands" (Coleman 2012a, 145). Also, if the Hogan twins exhibit overlap—whether of perspectives or contents—it is neither (presumably) necessary nor total overlap. Thus none of these examples, Coleman would presumably say, illustrates how subjects, or perspectives, or minds could enter into a state of connubial bliss. Minds can marry, on Coleman's view, only if there could be two subjects whose introspective access necessarily ranged over the same set of phenomenal elements. However, there could never be such a case, he avers, since "to say that there are two entails that the introspective access of one could differ as to some phenomenal element, with respect to the other" (Coleman 2012a, 145).

Yet perhaps Coleman is too eager to admit impediments to the marriage of true minds. First off, it is not entirely clear what mental contents are supposed to be when bereft of perspectives. Coleman does have something to say about this, but let me put it aside; for it seems that even if the notion makes sense, the Hogan twins, arguably, illustrate an example of shared perspectives. Of course, they do not illustrate necessarily shared total perspectives, but partial overlap of perspective would seem sufficient for pooling. When the blood of Jack the Ripper's victims pools under their necks, the drops do not overlap entirely giving you just one drop, but rather the drops cohere and we are left with a ghastly crimson pool. Since with cohesion we get only partial overlap at the edges, perhaps the partial overlap of perspectives is enough to at least help us make sense of minds or subjects combining at a level. And it would seem that even if the fantastical cases of partial phenomenal overlap via telepathy or split brain cases do not illustrate overlap of perspectives (and, in any event are too iffy to begin with), the case of the Hogan twins may give us a better sense of how this could be so.

8.5 A Refutation by Boredom?

Another reason Coleman has for thinking that the panpsychist ought not to posit subjects in the world's innermost core is that doing so is insufferably boring: "in fact," he tells us, "this sort of manoeuvre is boring in such a deeply metaphysical way that this alone indicates, from what we know of the workings of the world, that what we have on our hands is far from the correct solution" (Coleman 2012a, 149). This is an interesting counter to Ockham's razor, which tells us if two hypotheses are equally consistent with the data, we should prefer the simpler: rather than Ockham's razor, we have a complicating device, Coleman's curling iron, which tells us that if two hypotheses are equally consistent with the data, we should prefer the more complex. Employing this principle, we see that although explaining big subjects in terms of little subjects is, at least on some ways of measuring things, simpler than explaining subjects in terms of something else, its deathly dullness should lead us to reject the view. But, just as applications of Ockham's razor are countered by the observation that there is little evidence that simple theories are more likely true than more complex ones, there is little evidence that interesting theories are more likely true than boring ones (though perhaps they are more likely to get published).

I'm not persuaded by the suggestion that taking the ultimate constituents of the world to be subjects—or in other words, placing subjects at the fundamental level—is too boring to be true. However, I would say that it is a good methodological principle to try to explain things rather than taking them as fundamental. And in calling a theory of panpsychism that posits microsubjects "boring" (Coleman 2012a, 149), this is really what Coleman means. (Curling irons aren't in his tool box anyway—it's a good look, Sam, don't change it!) That is, Coleman is not claiming that taking subjects as fundamental is soporific, but rather that it leaves something unexplained that seems to call for an explanation: we should "want to say," he tells us, "something remotely interesting about how minds come about, not simply take them so thoroughly for granted" (Coleman 2012a, 149). This is a reasonable request: Why give up and simply call something fundamental rather than forging ahead on our quest to understand it in other terms? Yet, isn't this just what the panpsychist and Coleman himself are doing with subjectless experience, or what he calls "qualities"? As Coleman points out, although he doesn't posit fundamental subjects, he does posit fundamental experience (yawn), and thus is not furthering our quest to explain experience in terms of something else. Yet in doing so he sees himself as taking fundamental only what is minimally necessary. Well, maybe; but on the other hand, I tend to favor the view that one should never give upin philosophy or in life, for that matter.

8.6 Fundamental Continuity

Panpsychism is sometimes explained as the view that, in David Chalmers's words, "the fundamental physical entities [such as quarks and photons] have conscious experiences" (this volume). However, if our current physics is in the right ballpark, the fundamental nature of reality may not comprise discrete particles but rather the continuous *fields* of quantum field theory. As the physicist David Tong puts it, "[t]he objects we call fundamental particles are not fundamental," rather "they are ripples of continuous fields, moulded into

apparently discrete lumps of energy by that framework of quantum mechanics" (Tong 2012, 4). It is an open question whether this is true, yet if it is true, and if subjects are discrete (even if they are capable partial pooling), then it would seem that subjects are not part of the world's innermost core.

If Tong's picture of reality is correct, then it would seem that panpsychists should think of the fundamental nature of the world as comprising not discrete particles, but rather a continuous expanse of consciousness, what might be called "psych," which should be thought to refer not to the soul or spirit-like "psyche" of the panpsychists of yore, but rather the underlying experiential, nondiscrete nature of the universe. On this picture, God, as it were, did not make the elementary particles; rather, She made psych, and from that all else emerges. If our current physics is in the right ballpark, psych may correspond to the fields of quantum field theory. And fields, if our current physics is a guide, are both fundamental and continuous. Particles, then, for the panpsychist, are mere ripples in psych.

Since discrete particles emerge from fields, the panpsychist on this picture is still faced with explaining how conscious particles can combine to create human-level conscious experience. Yet, as I would like to now suggest, once we reject the pooling metaphor, combination is a relatively straightforward task.

8.7 Panpsychism in a Layered World

The Hogan twins' thalamic bridge (arguably) connects human-level mind to human-level mind. And this (arguably) partial perspectival overlap, I claimed, might suffice to show that subjects can pool. But does the panpsychist really need pooling in the first place? Or, rather, does the panpsychist need pooling to occur in the somewhat literal sense that Coleman uses the term, that is, pooling in the sense that smaller quantities of a thing come together to make larger quantities of that same kind of thing the way droplets of mercury pool in a mercury pendulum clock.

A molecule is a grouping of atoms, but atoms do not need to literally pool in order to create molecules; a macromolecule is a grouping of molecules, but molecules do not need to fully blend into each other to create a macromolecule; a cell is a grouping of macromolecules but the discrete nature of the lower-level molecules remains in the grouping, and so forth. Generally when we move from lower-level microstates to higher-level macrostates, the lowerlevel microstates retain their individuality to a degree. And Coleman readily admits this: "[i]n the combination of hydrogen and oxygen to form water, for example, . . . the constituent atoms *continued to exist*, albeit in modified form, after the integrated whole was formed" (Coleman 2012a, fn. 16). One oxygen atom and two hydrogen atoms considered at the atomic level are discrete individual entities, yet when brought together change to form water at a higher level; understood at a lower level, we have bonded atoms; at a higher level, we have water. There is no total overlap and each retains its identity to a degree in the bond; yet they do marry. Why shouldn't the same be true of minds? For the panpsychist, microminds exist as distinct entities at the microlevel, but at higher levels they undergo some changes and, while retaining their identity to a degree, enter into the everlasting bond of marriage.

Why then, can't it be that just as the nonreductive physicalist grounds consciousness in lower-level physical processes of the brain which in turn are grounded in lower—and still-lower-level processes until we reach the fundamental level (or if there is no fundamental level, then so on ad infinitum), that the panpsychist grounds human-level consciousness in lower-level microconsciousness? Of course, even with H_2O combining to form water, nature brings in very sophisticated engineers. And with the combination of subjects, the engineering problem may be even more difficult. But I fail to see why there should be a special problem—the combination problem—that is supposed to make panpsychism impossible. That is, I fail to see why combination should be or even appear to be intractable when our raw materials are conscious as opposed to when they are not.

Chalmers (this volume) argues that the biggest question for a view such as this is: "What is the phenomenal bonding relation?" This may be a big question and of great interest, however, I do not think that it is incumbent on the panpsychist philosopher to answer it; all that is required at this point in the debate, it would seem, is to make conceptual room for the position. And I see no conceptual impediment to the view that subjects, when combined in the right type of complex organization (whatever this may be), could ground a higher-level subject. Of course, it would not be simply that complexity accounts, or at least fully accounts for subjectivity. Rather, the raw materials are already conscious entities with perspectives and because they are already conscious perspectival entities, subjecthood comes for free. This, it would seem, is an advantage since it allows panpsychists to avoid the "structural argument" against physicalism, that is, the view that from structure and relations all one gets is more structure and relations and never anything like consciousness. And although panpsychists may face other deamons, such as providing an account of just what microsubjects could be, avoiding the structural argument is an important advantage.

But what of James's idea that human minds never combine into a corporate mind? Doesn't this at least indicate a difficulty with the combination of subjects at the lower level as well? One response is simply that since panpsychists need not say that every combination of minds creates a higher-level mind—for, as I said, some things combine, others don't—they might want to hold that microminds combine into human-level minds, but that human-level minds do not combine to form higher-level minds. However, it seems to me that the panpsychist need not even reject the idea of human minds combining to form higher-level minds. James and Coleman point out that humans don't pool merely by stacking. Thus when you place them one on top of the other, like a deck of cards, you don't get a group mind. I've noticed this as well. However, at a higher level, humans and social infrastructures interact in complicated ways that create communities and social organizations, and such conglomerates, though not human minds, could be thought of as group minds.⁴

Arguably, a group mind is created in the complex social interactions among bees. As the animal behaviorist Thomas Seeley sees it, "the process of evolution, operating over millions of years, has shaped the behavior of bees so that they coalesce into a single collective intelligence" (Seeley 2010, 19). Or as he puts it in more detail:

It [is] useful to compare what is known about the mechanisms of decision making in the bee swarms and primate brains. This may seem a bizarre comparison, for swarms and brains are vastly different biological systems whose subunits—bees and neurons—differ greatly. But these systems are also fundamentally similar in that both are cognitive entities that have been shaped by natural selection to be skilled at acquiring and processing information to make decisions. (Seeley 2010, 199)

To be sure, just because some believe, or perhaps simply just write as if they believe, that a colony of bees can form "swarm intelligence" or a "hive mind" does not indicate that there could be such a thing.⁵ However, it should weaken the conviction that the social mind is conceptually impossible. Regardless of whether there is such a thing as a hive mind, it is not obvious that the idea is unintelligible. Indeed, there seems to be no conceptual barrier to thinking that the highly complicated social organization of aliens in spaceships that Block describes in his thought experiment do combine to form a mind; they may not form a human mind, but perhaps they do form a social mind. Such considerations suggests that the purported conceptual difficulty of combining minds, might not be so conceptually difficult for everyone.

An advocate of the view that panpsychism is threatened because macrominds do not combine might respond to this by claiming that is the conceptual difficulty of combining not "intelligences" but perspectives or phenomenology that is supposed to be the stumbling block for the panpsychist. Nonetheless, the idea of a "group mind" does speak to the basic objection, voiced by James, Coleman, and others, that no matter how you arrange them, higher-level minds do not combine; in terms of this being conceptually impossible, some seem to find no problem with the idea of minds combining.

Of course, one may be dissuaded from accepting the view that social organizations can be conscious because one fears the incredulous stare. However, if the incredulous stare does not prevent the panpsychist from admitting microminds into her ontology, why should it make her demure when it comes to admitting what James disparaged as the "spirit of the age," and the "sentiment of the people"? If you are a panpsychist, it would seem that such talk need not be metaphorical.

Tolstoy, in *The Kreutzer Sonata*, tells of another way in which minds may merge. In the story, which recounts a deed more horrifying than those of Jack the Ripper, the central character speaks of music carrying him "instantly and directly" into its composer's mind: "my soul merges with his, and together with him I'm transported from one state of consciousness into another" Tolstoy (1889/2008, 96). If panpsychism is true, we can come closer to understanding the possibility of such a merger.

8.8 Why Panpsychism in the First Place?

Still, even if there are no barriers to combining minds, it might not seem that combinations of microminds could necessitate a macromind. And if so, panpsychism might not be worse, but would still be no better off than physicalism at bridging the explanatory gap. This is exactly what Goff (2009) claims. In brief, his argument is this: Imagine that panpsychism is true and that the fundamental nature of our world is imbued with mentality. Yet, he tells us, we can coherently imagine that the fundamental properties of this world are duplicated without human mentality being duplicated. That is, we can imagine panpsychist zombies. And assuming that coherent conceivability of a situation implies the possibility of that situation (as he thinks that panpsychists are wont to do), this means it is possible for there to be a panpsychist world such that duplication of its microlevel properties fails to duplicate human-level consciousness. Thus panpsychist microminds do not necessitate macrominds and thus panpsychism has no advantage over physicalism in solving the mind-body problem. Is this right?

One might, of course, reject the possibility of panpsychist zombies by rejecting the implication from coherent conceivability to possibility. Yet, as Goff points out, if one is willing to do that, one might as well (given how he sees the other checks and balances lining up), accept physicalism; as he puts it, "there does not seem to be a relevant difference between the two cases [panpsychism and physicalism] which could justify advocating the move from conceivability to possibility in the one case but not the other" (Goff 2009, 309). ⁶ I'm inclined to reject the idea that coherent conceivability entails possibility (unless, of course, coherent conceivability is defined as possibility, as, from what I can tell, it sometimes implicitly is) and to seriously question whether we have a rich enough conception of the microfeatures of the world to get the thought experiment started in the first place. However, given that these inclinations can be used in defense of physicalism just as readily as panpsychism, the question again arises: Why consider panpsychism in the first place?

I think the reason is not that one can apply the zombie argument in one case but not the other, but rather that the insertion of mentality at the fundamental level makes it easier to see how the fundamental level could necessitate consciousness. And this, I take it, though it doesn't tip the scale, is a point in the panpsychist's favor.

For James, who held that nothing exists save for "the everlasting atoms" and that higher-level things have "no existence out of our mind," it was reasonable to conclude that everlasting mental atoms would also not combine (James 1890/1950, 161). However, once we give up the idea that discrete things never combine, there seems to be no good reason to think that the combination problem exists at all.

Is panpsychism a better solution to the mind-body problem than physicalism? The answer to this question depends on what virtues and vices (apart from the combination problem, if my dissolution of it has been successful) one finds in each. Chalmers weighs each position carefully and finds that some type of antiphysicalist position tips the balance. I think there may be a heavy hand on that scale; however, this is not the place to lay out all of my reasons for why.⁷

Notes

- 1. This may be true, as Kuhn (1996) saw it, in the sciences as well, but I see it as especially pertinent to philosophical inquiry since what is philosophy, if not exciting?
- 2. He refers to himself in this manner in Coleman 2012b.
- 3. Would each hemisphere in split brain cases have introspective access to the other? And if they do, would we still count each hemisphere as grounding a distinct subject? It is difficult to say.
- 4. For a discussion of the social mind, see Bryce Huebner (2014). And see Wimsatt (1994) for a discussion of the relation between higher and lower levels of organization.
- 5. Cf. Zimmer (2012).
- 6. One might also argue along lines I suggest in Montero (2013 and *forthcoming*) that panpsychism does not require the supervenience of the macromental on the micromental. But such considerations also apply equally well to both cases.
- 7. Where is the place? My numerous papers, of course, that are still unwritten! Yet see Montero (2003; 2005; 2006; 2007; 2010) for an indication of how I see some of the costs and benefits.

Bibliography

- Block, Ned. "Troubles with Functionalism." *Readings in the Philosophy of Psychology*. Ed. Ned Block. Vol. 1. Cambridge, MA: Harvard University Press, 1980. 268–306.
- Coleman, Sam. "Mental Chemistry: Combination for Panpsychists." *Dialectica* 66.1 (2012a): 137–66.
- Coleman, Sam. "Review of 'The Mental as Fundamental: New Perspectives on Panpsychism.'" Notre Dame Philosophical Reviews. 17 Jan. 2012b. Web. 14 July 2014.
- Dominus, Susan. "Could Conjoined Twins Share a Mind?" *New York Times Magazine*. 29 May 2011. Web. 14 July 2014.
- Goff, Philip. "Experiences Don't Sum." Journal of Consciousness Studies 13.10–11 (2006): 53–61.
- Goff, Philip. "Why Panpsychism Doesn't Help Us Explain Consciousness." *Dialectica* 63.3 (2009): 289–311.
- Huebner, Bryce. Macrocognition: A Theory of Distributed Minds Collective Intentionality. Oxford: Oxford UP, 2013.
- James, William. The Principles of Psychology. Vols. 1, 1890. New York: Dover, 1950.
- Kuhn, Thomas. The Structure of Scientific Revolutions. Chicago: U of Chicago P, 1996.
- Montero, Barbara. "The Epistemic/Ontic Divide," Philosophy and Phenomenological Research 66.2 (2003): 404-18.
- Montero, Barbara. "What Does the Conservation of Energy Have to Do with Physicalism?" Dialectica 60.4 (2006): 383-96.
- Montero, Barbara. "Physicalism Could Be True Even If Mary Learns Something New." Philosophical Quarterly 57 (2007): 176–89.
- Montero, Barbara. "A Russellian Response to the Structural Argument Against Physicalism." Journal of Consciousness Studies 17.3-4 (2010): 70-83.
- Montero, Barbara. "Must Physicalism Imply the Supervenience of the Mental on the Physical?" Journal of Philosophy 110.2 (2013): 93–110.
- Montero, Barbara. "Irreverent Physicalism." Philosophical Topics (forthcoming).
- Montero, Barbara, and Papineau David. "The *Via Negativa* Argument for Physicalism." *Analysis* 65.3 (2005): 233–37.
- Seeley, Thomas. Honeybee Democracy. Princeton, NJ: Princeton University Press, 2010.
- Shani, Itay. "Mind Stuffed with Red Herrings: Why William James' Critique of the Mind-Stuff Theory Does Not Substantiate a Combination Problem for Panpsychism." Acta Analytica 25.4 (2010): 413–34.
- Stoljar, Daniel. Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness. Oxford: Oxford UP, 2006.
- Tolstoy, Leo. The Kreutzer Sonata. 1889. London: Penguin Classics, 2008.
- Tong, David. "Physics and the Integers" FQXi Essay Contest 2011. Web. 14 July 2014.
- Wimsatt, William. "The Ontology of Complex Systems: Levels of Organization, Perspectives, and Causal Thickets." *Canadian Journal of Philosophy supp. Vol. 20*. Ed. Mohan Matthen and Robert Ware. Calgary: U of Calgary P, 1994. 207–74.
- Zimmer, Carl. "The Secret Life of Bees: The World's Leading Expert on Bee Behavior Discovers the Secrets of Decision-Making in a Swarm." *Smithsonian Magazine* March 2012. Web. 14 July 2014.

9

Panpsychist Infusion

WILLIAM SEAGER

9.1 Why Panpsychism?

Panpsychism is the view that consciousness is a simple, fundamental, and pervasive perhaps ubiquitous element of reality. I focus on consciousness because that is the mental aspect of the world that has most persistently resisted assimilation into the scientific picture of the world. To say that consciousness is fundamental is, essentially, to endorse the claim that consciousness cannot be thus assimilated. To say that it is simple is to associate it with the most basic constituents of reality.

Of course, many philosophers, as well as many thinkers from other disciplines who have devoted themselves to the problem of consciousness, would reject panpsychism out of hand. Some of them regard panpsychism as unworthy of a reasoned examination. John Searle (1997) regards the view as simply "absurd"; Colin McGinn (1999) as "ludicrous." Obviously, the idea that everything in the world, from quarks to nations, enjoys a conscious mental life similar to our own has little to recommend it. But throughout its long history very few, if any, versions of panpsychism have entertained such an outrageous extension of mind into nature. Panpsychists have instead postulated that consciousness comes in a vast range of forms that begins with something unimaginably simple, but still phenomenal in nature, and proceeds through more complex forms up to and perhaps exceeding the teeming, dynamic, and self-aware consciousness with which we human beings are familiar. And panpsychists have generally denied that all complex physical entities enjoy complex consciousness or even any consciousness at all.

Those of a physicalist persuasion will find the assignment of even simple forms of consciousness to 'simple' parts of nature repugnant to naturalist sentiments. Against this, panpsychists can offer two lines of reply, one negative and one positive. Since one of the primary reasons for rejecting panpsychism is its incompatibility with physicalism, the negative reply stems from the now long-standing failure of physicalists to integrate consciousness into the otherwise spectacularly successful, comprehensive, and ever-expanding scientific picture of the world. While it always remains possible to insist that consciousness will take a standard place in the physicalist view, it is undeniable that consciousness seems uniquely difficult to 'physicalize.'

It is a testament to the power of the antiphysicalist arguments that many physicalists have conceded the apparent arbitrariness in the link between mind and matter.¹ Their defense involves explaining why the evident 'explanatory gap' (see Levine 1983) is a *mere* appearance that does not license ontological implications. According to this strategy, there are two 'ways of knowing' consciousness, one as it were from the outside and one from the inside. The inside view's knowledge is couched in what are called phenomenal concepts. The knowledge available from the outside is expressed in ordinary physical concepts whose referent is whatever physical state is to be identified with consciousness (for attempts along these lines, see Loar 1990; Rey 1993; Papineau 2006). It is some difference between these two classes of concepts which explains why it is so hard to see how they refer to the same physical states.

This strategy faces a deeply problematic dilemma. Either it endorses the claim that phenomenal concepts possess Fregean senses or that they do not. What I mean by the first horn of the dilemma is simply that when we are conscious, we are presented with some features by which we distinguish one phenomenal state from another. These features then stand as new properties that have to be integrated into the physicalist account but that face exactly the same problem with which we began. In place of a single explanatory gap, there will be an open-ended hierarchy of gaps, the last of which will generate the same antiphysicalist worries as the first. This kind of objection has been called by Stephen White the "curse of the qualia," and he has developed it into an extremely powerful attack on the phenomenal concepts strategy (see White 1986, 2010).

The other horn of the dilemma is no less tractable. It depends on the claim that phenomenal concepts are bare or brute recognitional concepts that lack any substantive sense and thus do not depend on appreciation of any feature of experience for their application.

Although, in general, the existence of such bare recognitional concepts is not implausible, the problem with this approach is nonetheless pretty clear. It is wildly implausible that when we apply phenomenal concepts we do so in the absence of any 'source material' in experience on the basis of which we categorize phenomenal consciousness. Or, to put it another way, if application of phenomenal concepts was via such pure recognitional capacities, then this would be evident to us. Compare how you know how your limbs are currently arranged (without looking!) with how you know what colors you are experiencing. I know both, but the former knowledge does not seem to be mediated by any particular quality of my experience (save when my limbs are in unusual and uncomfortable positions or have been motionless for a long enough to generate pain), but my awareness of colors is obviously vividly phenomenological. The psychological literature is replete with examples of neurological disorders that feature what might be called knowledge without awareness, as in blindsight, but there are many others.² It is, of course, striking that what is missing in these cases is specific sorts of consciousness despite the presence of certain recognitional capacities.

If one takes the recognitional capacities approach to its logical conclusion, consciousness becomes a kind of illusion. On this view, there is no phenomenal experience, but we possess a rich and complex set of concepts that describe a nonexistent world in a proprietary manner.³ Recognitional capacities trigger the application of these concepts, and discursive thought over the long span of human cognitive development has elaborated them into a structure that supports a rich but delusive system of beliefs. In terms of what we think consciousness is *within* this system, we are actually no more conscious than rocks. Daniel Dennett provides a clear expression of this view, which involves

a neutral method for investigating and describing phenomenology. It 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)

I somewhat hesitate to ascribe this view to Dennett since his writing is often ambiguous between a position devoted to debunking certain perhaps dubious philosophical notions, such as that of qualia, and a position which entails the wholesale denial that there is anything even remotely like phenomenal consciousness. The former attacks a straw man. The latter position is surely absurd. The problem of consciousness does not revolve around descriptions of consciousness but around the simple fact that conscious beings are *presented* with the world, and themselves, in a special way quite different from the causal and information-laden reactions of more ordinary physical objects.

The idea that this sort of 'presence' is a fictional object seems too wildly implausible to be taken seriously, yet seems the natural upshot of the pure recognitional capacities interpretation of phenomenal concepts. Thus the negative argument for panpsychism seems genuinely powerful. Physicalism is far from established and cannot be deployed in an instantaneous and unassailable refutation of panpsychism.

However, stepping back from physicalism does not entail panpsychism. The most important positive argument for panpsychism was put into its canonical form by Thomas Nagel (1974/1979). Nagel's argument can be summarized thus:

- P1. Consciousness is either a fundamental or an emergent feature.
- P2. Consciousness is not an emergent feature.
- C. Therefore, consciousness is a fundamental feature.

Why does Nagel think that consciousness cannot be an emergent feature? This question is pressing because, as we shall see, consciousness will turn out to be an emergent even if we endorse panpsychism. Despite this, Nagel's argument remains cogent. How can this be?

The key is a distinction between what I will call conservative versus radical emergence. Nagel's claim in P2 is shorthand for the dual claims that consciousness is not a conservative emergent from a purely physical submergent base and that radical emergence is impossible.

Roughly speaking, a conservatively emergent property of an object O is one whose exemplification by an object follows logically from the specification of the properties of O's constituents (plus environment) and the laws governing these submergent properties. Thus, the liquidity of water (at standard temperatures and pressures) is a conservative emergent feature because while neither oxygen nor hydrogen atoms exemplify liquidity, it follows logically from the properties of oxygen and hydrogen and the physical laws which govern their interactions, that large samples of H_2O must form a liquid under ordinary conditions. It is absolutely impossible for oxygen, hydrogen, and the environment to be as they are, and the laws of physics to be as they are, and for large samples of H_2O to fail to be a liquid.

By contrast, radical emergence involves a weaker link between the submergent base and the emergent properties generated by it. Radically emergent properties are still lawfully dependent on the distribution of properties over the submergent entities, but some of the laws are special laws of emergence that must be appealed to in addition to the laws of the submergent domain. These laws link the submergent to the emergent features and can vary across possible worlds. This sort of law was referred to as "trans-ordinal" by C. D. Broad, who explicitly evoked the image of levels of reality. In these terms, there are laws restricted to a single level: the purely physical laws, the laws of chemistry, the laws of biology, and so forth. Transordinal laws generated radically emergent properties out of the features at the submergent level(s). For example, Broad believed that chemical properties were not strictly necessitated by the purely physical level and sprang into being when the physical level attained the appropriate state via the action of transordinal laws of chemical emergence (see Broad 1925, ch. 2). But Nagel denies that such laws of emergence are possible because

there are no truly emergent properties of complex systems. All properties of complex systems that are not relations between it and something else derive from the properties of its constituents and their effects on each other when so combined. (Nagel 1974/1979, 182)

Nagel is suppressing the distinction between radical and conservative emergence here. There doesn't seem to be any logical ground for denying the possibility of transordinal laws.⁴

Let us, however, concede for the moment that radical emergence is off the table. Then the cogency of Nagel's argument depends on the claim that consciousness cannot be a conservative emergent feature stemming from purely physical submergent features. Of course, the core antiphysicalist arguments alluded to above all tend in exactly this direction. They all tend to show that there is no way to get consciousness out of the physical in anything like the standard way. If that is correct then consciousness cannot be an emergent. If consciousness is not emergent then it is a fundamental feature of the world.

It is but a small step from the idea that consciousness is fundamental to the core tenet of panpsychism: Consciousness is a characteristic of the most simple elements of nature. At least, this follows if we accept the general scientific account of the long-term development of the world, which begins with a rather undifferentiated 'sea' of very simple things gradually developing into more and more complex entities interacting according to the dictates of physical law. Otherwise, we have the inexplicable appearance of a fundamental feature of the world—consciousness—at some late date in the universe as a property of certain highly complex entities. Such a scenario is just radical emergence that we are assuming does not occur. Thus consciousness in its simplest or most basic form should be associated with the simplest and most basic physical features that constitute the universe.⁵ To this extent, Nagel's argument for panpsychism is vindicated.

However, once we relax the stricture against radical emergence, we face a disjunctive conclusion: either panpsychism is true or else there is radical emergence. I don't see any strong proof that the idea of radical emergence is incoherent, but that does not mean it is attractive. In fact, there are a number of basic considerations that disfavor it.

For one thing, it is metaphysically uneconomical. If we can do without positing *both* new fundamental features and peculiar cross-domain laws governing their creation ex nihilo then we should avoid it.

For another, there seems to be only one place in the world where radical emergence looks to be even potentially necessary and that is the case of consciousness. This is at best a strange intrusion into an otherwise well-behaved world. In the absence of any other clear case of radical emergence this also makes its postulation in the unique case of consciousness seem ad hoc.

For yet another, radical emergents standardly herald the arrival of new causal powers into the world. Yet the structure of the physical world seems to be complete and causally closed.

Panpsychism holds out the promise of a different picture. One in which consciousness is a fundamental feature of the world, irreducible to the purely physical but one that otherwise fits quite smoothly into the extant, extremely attractive and successful, scientific picture of the world.

9.2 Deferential Emergentist Panpsychism

The version of panpsychism that best exemplifies these advantages I call deferential emergentist panpsychism (DEP). The view is deferential in the sense that it accepts that modern science provides an accurate picture of the world, which is entirely physical and exhibits a hierarchical structure of increasingly complex emergent properties and systems. The kind of emergence required to fund the scientific picture is restricted to conservative emergence. DEP is emergentist in this minimal sense. But more than the standard conservative emergence that typifies the physical world, DEP also postulates that consciousness forms its own system of emergence, which is similarly conservative, and reflects in some way the growth of mentality in correlation with the increasing complexity of physical systems.

DEP presents the world as exemplifying consciousness from its inception, as nonphysical properties of the fundamental physical entities that constitute the world. As these entities interact to form more complex physical entities, so too does consciousness become more complex. Presumably, the sort of consciousness that the physical fundamentals enjoy is of a simplicity which we can scarcely imagine. Furthermore, we have little grasp of the laws, or the "mental chemistry" (Nagel 1974/1979, 182), by which more complex states of consciousness emerge.

This phrase was not invented by Nagel but can be traced to John Stuart Mill (Mill 1869/1989, 108–9). Mill was no panpsychist; he was, in our terms, a radical emergentist.⁶ His mental chemistry expanded the range of associationism

in psychology to include combinations of sensations into new forms (Mill uses the analogies of the spectral composition of white light and the way a moving light forms a visible 'trail'). For Mill, mental chemistry, as indeed physical chemistry, involved a kind of radical emergence. But it is interesting that the explicit idea of combining mental states into new forms can be found in Mill's writing. Despite his emergentism, he makes a number of pronouncements about mental chemistry that will figure in the articulation of DEP. For example, Mill continues the discussion of the way mental composition occurs, extending the analogy with the spectrum:

If anything similar to this obtains in our consciousness generally . . . it will follow that whenever the organic modifications of our nervous fibres succeed one another at an interval shorter than the duration of the sensations . . . those sensations or feelings will, so to speak, overlap one another, and becoming simultaneous instead of successive, will blend into a state of feeling, probably as unlike the elements out of which it is engendered, as the colour white is unlike the prismatic colours. (Mill 1869/1989, 108)

Mill also considers how mental phenomena might join to create new forms of consciousness in his *Logic*:

The generation of one class of mental phenomena from another ... is a highly interesting fact in psychological chemistry; but it no more supersedes the necessity of an experimental study of the generated phenomenon than a knowledge of the properties of oxygen and sulphur enables us to deduce those of sulphuric acid without specific observation and experiment. (Mill 1843/1963, 534)

And further:

[I]t appears to me that the Complex Idea, formed by the blending together of several simpler ones, should, when it really appears simple ... be said to result from, or be generated by, the simple ideas, not to consist of them. (Mill 1843/1963, 533)

These remarks reveal Mill's commitment to what we're calling radical emergence and also Mill's sense that it is impossible to 'compose' mental entities together in a purely constitutive manner. Rather, when the right mental entities occur together or in the appropriate relation, then they will 'generate' a novel form of consciousness. It is thus a crucial question for panpsychism how to understand 'mental chemistry.' DEP must appeal to something like mental chemistry, but the issue remains whether this can be understood as requiring nothing more than conservative emergence. If radical emergence must be invoked to make the leap from the foundational elements of microconsciousness to more complex forms then DEP becomes otiose. For, if radical emergence is accepted it is obviously simpler and more elegant to let complex consciousness radically emerge from purely physical fundamentals and forego the intuitively implausible assignment of mental properties to basic physical entities.

But radical emergence seems implausible in general and unmotivated in the purely physical domain. So there is hope that finding the right sort of mental chemistry that demands no more than conservative emergence will make DEP the more attractive position. However, one monster problem immediately raises its threatening and ugly head, a problem that may destroy any hope of finding anything like the mental chemistry we need.

9.3 The Combination Problem

As is now widely recognized, the combination problem presents perhaps the most difficult implementational problem for panpsychists (see Chalmers this volume). The classic statement of the problem is by William James (1890, ch. 6). James's basic complaint is that there is simply no way to make sense of the notion of mental combination. It is important to note that James is essentially taking over Mill's understanding of combination and constituency here. On this understanding combination operates by the straightforward summation of the combining elements that retain their identity and causal powers before, during, and after composition. This is what Mill called the mechanical mode of the composition of causes against which he opposed the chemical mode (see Mill 1843/1963, vol. 3, ch. 6). The distinction between the mechanical and chemical modes of causal composition corresponds to the distinction between conservative and radical emergence.

James's sense that mental combination in the mechanical mode is hopeless is clearly expressed in this passage:

Where the elemental units are supposed to be feelings, the case is in no wise altered. Take a hundred of them, shuffle them and pack them as close together as you can (whatever that may mean); still each remains the same feeling it always was, shut in its own skin, windowless, ignorant of what the other feelings are and mean. There would be a hundred-and-first feeling there, if, when a group or series of such feelings were set up, a consciousness belonging to the group as such should emerge. And this 101st feeling would be a totally new fact; the 100 original feelings might, by a curious physical law, be a signal for its creation, when they came together; but they would have no substantial identity with it, nor it with them, and one could never deduce the one from the others, or (in any intelligible sense) say that they evolved it. (James 1890, 160)

This problem has long bedeviled panpsychist thought. Almost two decades after James wrote the *Principles*, he was still banging his head against the combination problem:

I struggled with the problem for years, covering hundreds of sheets of paper with notes and memoranda and discussions with myself over the difficulty. How can many consciousnesses be at the same time one consciousness? . . . The theory of combination, I was forced to conclude, is thus untenable, being both logically nonsensical and practically unnecessary. (James 1909, 207)

James solution, if it is a solution, is a radical rejection of the terms of the problem. He regards the problem as a symptom of 'intellectualism' by which he means something like the forced deployment of concepts in domains where they are desperately and irredeemably inadequate. Perhaps despairingly, perhaps in liberation, he writes: "I have finally found myself compelled to give up the logic, fairly, squarely, and irrevocably. . .. I prefer bluntly to call reality if not irrational then at least non-rational in its constitution" (James 1909, 213).

The combination problem also confounded the most famous panpsychist of the twentieth century, Alfred North Whitehead. He castigates Leibniz for failing to solve the problem that Whitehead claims is "a perplexity which is inherent in all monadic cosmologies" (Whitehead 1929/1969, 32). Recall Leibniz's example of our perception of the sound of a wave crashing on a beach, which perception is literally composed of untold numbers of unconscious auditory perceptions of the individual droplets hitting the beach. The resultant consciousness, in which the multitude has been lost, Leibniz calls "confused." Thus, we might say, confusion solves the combination problem for Leibniz. But without an analysis of how confusion arises this is a merely verbal solution. As Whitehead says, "he [Leibniz] fails to make clear how 'confusion' originates" (Whitehead 1929/1969, 32). Whitehead describes the problem as follows:

[T]he integration of simple physical feelings into a complex physical feeling only provides for the various actual entities of the nexus being

felt as separate entities requiring each other. We have to account for the substitution of the one nexus in place of its components." (Whitehead 1929/1969, 293)

This process Whitehead calls "transmutation." It is very complex, involves a great deal of Whiteheadian machinery, and I will not pretend to understand it. Notice that it is a multistep process, for the initial transmutation takes a set of 'physical feelings' being taken up into an initial unity by which each is felt to 'require' the others. This is analogous to James's one hundred feelings being partially integrated. A second stage of transmutation goes further and somehow or in some way 'absorbs' the feelings into a single unitary state.

Nonetheless, we can take these pronouncements of our great forebears Leibniz ('unity by confusion'), James ('unity by generation'), and Whitehead ('unity by transmutation') as clues to the outline of a solution to the combination problem.

9.4 Infusion Not Combination

The solution I will advance here I call 'combinatorial infusion.' It follows James's suggestion (which he found inadequate) that the transmutation from the hypothetical micropsychic features assigned to the fundamental entities of the physical world to the macrostates of consciousness with which we are introspectively familiar requires the generation of a new state which infuses its precursors, or, to use Whitehead's term, substitutes a new state for the set of precursor states.

The two core constraints on combinatorial infusion are, first, that infusion be a form of conservative emergence and, second, that it provide some kind of intelligible link between the micropsychic features assigned to fundamental physical entities and resultant complex states of consciousness. Given our concession to James's denial that mental combination is possible and that only radical emergence of consciousness is possible, it looks very difficult to meet these constraints.

The difficulty arises, however, from the rather limited view of combination that James endorses. James implicitly identifies combination with what Mill called the mechanical mode of causal composition. Perhaps he agreed with Mill that any alternative would have to involve a kind of radical emergence. But James (and Mill) must be wrong, because we have examples from modern science that transcend the mechanical mode of causal composition but which do not invoke radical emergence. Let me briefly note two examples.

The first stems from quantum mechanics. Entanglement can create states that, at least in some cases, result in essentially new systems which have properties distinct from those of their precursors and causal powers which are not purely mechanical or additive results of the causal powers of their components. A specific example can be drawn from Paul Humphreys who regards quantum entanglement as involving a kind of fusing of entities into new systems.⁷ In the standard example of entanglement, the so-called singlet state, two particles interact so as to form a new state whose mathematical representation cannot be decomposed into a product of the representations of the constituents. The system acts as a unified state insofar as measurements on one part instantaneously put constraints on measurement results on other parts of the system and there is no way to determine whether the particles are entangled by any local measurement performed on the parts.

The second example is from classical physics. The famous 'no hair' conjecture about black holes states that they can be exhaustively characterized by three physical properties: mass, electric charge, and angular momentum. That is, no matter how a black hole is formed it will be indistinguishable from any other that shares the same mass, charge, and spin. In a certain sense, a black hole forms something like a fundamental particle. The physical entities that form a black hole can be said to fuse into a new entity which cannot be understood as a relational structure of its precursor entities. They have gone out of existence. The new system retains certain physical properties even as it throws away the particular characteristics of the precursor entities.

It must be emphasized that this example is simply used as a model for combinatorial infusion. The classical black hole is a product of general relativity and classical electromagnetism, and we have every reason to doubt that the complete physical story will leave its description unchanged. Be that as it may, the classical black hole stands as a viable model for the kind of fusion I am trying to articulate and apply to the combination problem.

These examples, within their own domains, meet analogues of the constraints laid down above. The kind of emergence they exemplify is *not* radical. This is evident from the fact that entanglement and black hole formation are predictions of the applicable physical theories and from the fact that the features of the emergents are drawn from the set of fundamental properties of the theories which predict them.

Panpsychist combinatorial infusion would thus postulate a set of fundamental properties of consciousness that are assigned to the fundamental physical entities which constitute our world. It holds that, under certain conditions of which we remain quite ignorant, mental fusion will occur, generating a new state of consciousness that is a function of the precursor states. This fusion would be a psychological process; it would not be the fusion of physical entities into a state of consciousness (a process of doubtful coherence and certainly a kind of radical emergence). It would be a fusion of mental entities into a new fused mental state. I do not think that fusion would be a hierarchical process. Although more complex physical states of the right sort would be the physical signature of mental fusion, the to-be-fused entities would be the fundamental mental states of the basic physical constituents of these complex physical states. For example, we know something about the neural correlates of consciousness. These highly complex states of the brain are aspects of a biological organ which has a multifaceted interactive and hierarchical functional architecture. But the mental fusion that such brain states occasion is the fusion of the fundamental micropsychic features⁸.

A natural question to ask at this point is what, exactly, are the fundamental micropsychic features? No very informative answer can be given at this point. DEP, based upon combinatorial infusion, is committed to some 'primitive' mental features but their specification is hostage to future theories of consciousness. For example, suppose for the sake of argument that the recent AIR theory of consciousness (see Prinz 2012) is correct, modulo the need to embrace DEP in order to introduce consciousness into the world.⁹ Prinz holds that mechanisms of attention directed on intermediate-level sensory representations engender consciousness. He maintains that all consciousness is sensory consciousness, construed broadly enough that the somatic and proprioceptive states of bodily awareness count as sensory. Prinz also puts tremendous effort into specifying likely neural correlates of sensory states and attention. In terms of a combinatorial infusion based DEP, these neural correlates point to the physical configurations that occasion mental fusion. Further, Prinz's claim that all consciousness is sensory suggests that the micropsychic features we need to postulate are elementary forms of such consciousness. Fused states will then arise that conjoin and blend these into more complex sensory states (such as phenomenal color, auditory, olfactory experiences, and the various bodily experiences of pain, pleasure, heat, cold, etc.).

On the other hand, many find the restriction of consciousness to sensory states implausible. If so, the range of elementary micropsychic features would have to be expanded to generate the states which the correct theory of the phenomenological nature of consciousness decrees. The main point is that while it seems hopelessly mysterious to posit that phenomenal experience is generated by a system of entities possessed only of physical properties such as mass, spin, charge, energy and a few more abstruse quantum mechanical properties, it does not seem similarly paradoxical that a system of entities possessed of elementary phenomenal features could generate complex states of consciousness, given an operation like combinatorial infusion whose intelligibility is clear and established in other domains.

Prinz's discussion of the neural correlates raises another important issue. This is the question of whether the physical states that occasion mental fusion must themselves exhibit their own form of combinatorial infusion. I think it would be fair to say that, broadly speaking, the various quantum theories of consciousness (see e.g., Hameroff and Penrose 1996; Hameroff and Powell 2009; Stapp 1993; Lockwood 1989) suggest an affirmative answer insofar as they expect to find in the brain some quantum 'signature' of consciousness that lines up quite well with the fusion-like features of quantum mechanics. While this is a real possibility and is ultimately an empirical question, I don't think that combinatorial infusion-based DEP needs to take a stand on this issue. It seems perfectly possible that 'ordinary' brain states, such as the ones Prinz selects as the correlates of consciousness, could be the physical state that underpins mental fusion. This is a good thing because there is little evidence that distinctively quantum effects play a role in the brain's functioning or that the brain cannot be understood, under the broad umbrella of the neuron doctrine, in standard neural network terms. In any case, the discovery of distinctive quantum effects in the brain would not undercut the need for combinatorial *mental* infusion. Quantum properties, even if undergoing something like fusion themselves, do not make the emergence of consciousness intelligible. If we posited, to take a ridiculous example that makes the general point, that when two electrons become entangled an experience of red appeared this would be just as mysterious as any other physical basis for the appearance of consciousness.

The best way to understand the place of consciousness in the world from the point of view of a combinatorial infusion-based DEP is something like a dual-aspect theory. That is, it would be a misunderstanding to think that the elementary micropsychic features are simply 'extra' properties of things which stand in the same relation to, say, electric charge as does, say, mass. It is better to think of consciousness as the expression, in the realm of the mental, of the kind of physical complexity that occasions mental fusion. On this view, the physical arrangements that matter to consciousness are necessary preconditions for mental fusion because they, as it were, 'line up' the micropsychic features in the right way. The best way to investigate these conditions is standard research into how consciousness is realized in the brain. This is another aspect of the deferentiality built into the account that minimizes disruption of the scientific picture in general and our developing physical understanding of the relation between mind and brain in particular.

9.5 Objections

David Chalmers (this volume) develops a characteristically clear and comprehensive overview that presents a set of core challenges for panpsychism raised by the combination problem. He also makes some criticisms specifically directed at a combinatorial infusion-based DEP. It is tremendously helpful to have such a focused set of issues at hand, and I conclude with some responses.

The Aggregation Argument. Stemming from some remarks of William James, this objection depends on the claim that aggregates are not objectively real. If we add the claim that all conservative emergents are aggregates, then we can conclude that consciousness (or a conscious subject) is a mere aggregate and hence unreal. Chalmers notes that this puts consciousness on a par with ordinary objects like rocks, tables, trees, and so on, and thus does not think the argument in this form is particularly worrying. I take the argument more seriously and think there is a real problem here (see Seager 2012b, ch. 9). But combinatorial infusion undercuts the worry. Infusions are not what James meant by an aggregate. They would instead fall under his remark quoted above that the constituent micropsychic features would "by a curious physical law, be a signal for its [the infusion's] creation" (James 1890, 160). What is distinctive about DEP is that it can embrace this remark *without* embracing radical emergence.

Subject Summing. A related argument sees DEP as requiring that subjects have to somehow combine to solve the combination problem and subjects simply do not sum. Following some work of Philip Goff, Chalmers spells out this argument in modal terms, with the crucial premise as: "It is never the case that the existence of a number of subjects with certain experiences necessitates the existence of a distinct subject" (Chalmers this volume). It is crucial to note that this assumes something about the nature of a panpsychism that, like DEP, wants more complex states of consciousness to arise from more elementary forms. The assumption is that it must be absolutely necessary that, in this case, the new subject arise from the old. One must be very careful about how to interpret such claims. To avoid radical emergence, they must indeed be necessary. But the absolute necessity at issue is one that is relative to the fixing of relevant laws of nature. Compare the case of the liquidity of water. This arises from the properties of hydrogen and oxygen and the laws of physics. Relative to these laws, it is absolutely necessary that water be a liquid (under standard conditions). To put it another way, there are simply no possible worlds where the laws of physics hold, where oxygen and hydrogen chemically combine under standard conditions, and where this combination is not a liquid.

In the case of a combinatorial fusion-based DEP, the laws that are relevant have to include the laws which underlie mental fusion. Given these laws and the physical laws, there are simply no possible worlds where a physical system gets into a state like those in our brains which occasion consciousness, where the micropsychic features are arranged as they are in the actual world, and where a new subject does not arise via combinatorial mental infusion. So the premise above can be legitimately denied. One note about subjects here might be in order. By subject I mean something very thin. A subject of a mental state is simply an entity that exemplifies the mental property which constitutes that mental state. When mental fusion occurs, subjects at issue are not the myriad elementary physical entities that bear the micropsychic features but rather the system that occasions the fusion (this might provide a new way to distinguish objectively real things from mere aggregates and solve the 'problem of the many').

The Conceivability Argument Against Panpsychism. Chalmers generalizes the subject-summing argument into one that mirrors the modal argument against physicalism. That argument works by endorsing the claim that the totality of physical truths and physical laws does not entail any fact about consciousness. The extended argument holds that the totality of physical truths, physical laws, and the micropsychic facts does not entail any macropsychic facts. So there is a possible world that is a physical and micropsychic duplicate of our world that is not a total duplicate or our world (in particular, this duplicate lacks complex states of consciousness). But, of course, the duplicate world should also duplicate any laws governing the micropsychic features. Combinatorial infusion-based DEP holds that there are such laws and that they (in concert with the arrangement of purely physical features) generate fused mental states. Naturally, if one does not include the micropsychic laws one can conceive of worlds that lack macropsychic features despite being otherwise identical to the actual world. By the same token, if one refused to include (some of) the physical laws one could conceive of worlds with oxygen and hydrogen but without water.

The laws that describe combinatorial mental infusion are presumably contingent. So there are indeed worlds that resemble our world with respect to the micropsychic features and the physical entities and laws but differ in their mental laws. That does not by itself seem to threaten combinatorial infusion-based DEP.

The Knowledge Argument Against Panpsychism. As above, this argument extends Frank Jackson's knowledge argument to the case of panpsychism. If we add to Mary's knowledge base all the micropsychic facts, she will still not be able to deduce the macropsychic facts. But again, of course not. She needs to know the laws of combinatorial infusion. It is no surprise that ignorance prevents deduction. Should we hold that the laws governing combinatorial infusion should be a priori knowable just from knowledge of what consciousness is like? I see no reason to think so. This means that combinatorial infusion-based DEP admits the possibility of zombies. True, but that is no problem. The *relevant* kind of zombies is absolutely impossible. The relevant zombies would be ones that occur in worlds that share our physical arrangements and laws, the micropsychic features' arrangements and the laws governing combinatorial infusion. There are no zombies in those worlds.
The Palette Argument. Read as an attack on combinatorial infusion-based DEP, this argument presents the worry that there will be too few micropsychic features to plausibly generate by infusion the myriad complex states of consciousness that can exist. Chalmers's own version of Russellian monistic panpsychism assigns micropsychic features in one-to-one correspondence with fundamental physical properties. There seem be relatively few of the latter, hence a paucity of the former.

One might reply by claiming that physically indistinguishable fundamental entities are nonetheless differentiated by their micropsychic features. However, this option faces the difficulty that since physical complexity is engendered strictly by physical law, the micropsychic character of entities is then irrelevant to the growth of physical complexity. In fact, this point was the basis of Nagel's argument for the ubiquity of fundamental micropsychic features.

Perhaps the best reply (as Chalmers implicitly notes) is to permit each kind of fundamental physical entity to host a constellation of fundamental micropsychic features which then combine according to combinatorial infusion as dictated by the laws of infusion and the variegated associated physical systems. This could perhaps be incorporated into a Russellian panpsychism by the claim that it is the *constellation* of micropsychic features associated with physically fundamental entities that provides the categorical basis for the dispositional properties revealed by scientific investigation. If we let our imagination take further flight, we could envision that since different constellations of micropsychic features make a difference in the kind of physical relations and causal interactions into which physical entities can enter, this in turn might affect the systems available to occasion mental fusion.

This reply entails, or at least strongly suggests, that each fundamental physical entity has a highly disunified consciousness in the sense that a number of incompatible basic micropsychic features belong to it. But disunified consciousness is not impossible, especially given the thin conception of subject that DEP endorses. In fact, combinatorial infusion suggests a way to think about the emergence of the unity of consciousness in which fusion inaugurates a kind of co-constituting totality in which each discernible feature is partly modified by and partly modifies its other components.

Finally, it is not completely obvious that a large number of micropsychic features is necessary to account for the range of macropsychic states of consciousness. After all, it takes but three fundamental hues to generate all possible colors that humans can experience. Of course, this is simply an analogy, but if, as noted in our discussion of Prinz's theory, all consciousness is sensory

consciousness, it might turn out that only a relatively small number of fundamental micropsychic features need to be postulated.

The Revelation Argument. This argument begins with the highly plausible claim that "[t]he nature of consciousness is revealed to us in introspection" (Chalmers this volume). Coupled with the further claim that "[i]f constitutive panpsychism is correct, consciousness is constituted by a vast array of microexperiences" (Chalmers this volume), we infer that this array of microexperiences should show up in introspection. Since it obviously does not, constitutive panpsychism finds itself in trouble. But it should be clear that combinatorial infusion-based DEP is in a different boat. Infusion generates a new entity that fuses the micropsychic in a way that erases its multiplicity. Think of the black hole example. Although formed by the gravitational collapse of a vast array of highly differentiated physical things, the resulting black hole keeps no record, so to speak, of its precursor 'constituents' and ends up exemplifying just a few new properties which are fusions of those of its initiating precursors. In introspection, we should expect to only find the fused macropsychic feature with no hint of the complex micropsychic features (and physical structure) needed to generate it. On the other hand, we might expect introspection to tell us something about the basic features themselves to the extent we can abstract away from complexity. We have, for example, found via introspection and investigation that color experience depends on only three basic experiential hues.

The Structural Mismatch Argument. I find this argument somewhat obscure, but the main point seems to be that if macropsychic states are constituted by micropsychic states that correspond to fundamental physical features, we should expect to see a match between macropsychic and macrophysical structure. Certainly, no such match is evident to introspection. Certain versions of what might be called purely constitutive or structural panpsychism could fall victim to this objection. But combinatorial infusion-based DEP will not because (1) the postfusion state will not reveal prefusion constituents, and (2) the physical state that occasions mental fusion need not have its structure duplicated in the fused mental state. The example of quantum entanglement might be useful here. There is no way to tell by looking at the entangled constituents of such a state whether they are entangled or not. Only by looking at the system as a whole will this be revealed. We can interpret this as a denial that the constituting structure must be duplicated or reflected in the resulting states.

Mental Causation. Perhaps the most serious problem for DEP, one that goes beyond the combination problem and bedevils wide swathes of the philosophy of mind, is how to ensure that mental features have causal efficacy.

Constitutive panpsychism attempts to solve this by postulating the micropsychic features as the categorical basis for the powers of fundamental physical entities. Then, "micro-phenomenal properties are causally efficacious in virtue of their playing fundamental microphysical roles, and macrophenomenal properties are causally efficacious in virtue of being grounded in microphenomenal properties" (Chalmers this volume). This issue raises many difficulties, but one response would be to appeal to the existence of fusion-like operations in the physical world. Entangled systems have powers distinct from those of their constituents taken by themselves. Perhaps fused mental states can stand as the categorical ground for fused physical states. This might suggest that we should look for such distinctive physical states as the correlates of states of consciousness.

On the other hand, a venerable view has it that we should regard the mental and physical as co-expressions of one underlying reality. In that case, we should not expect to find anything like psycho-physical causation. Combinatorial infusion-based DEP could then be deployed to explicate the nature of the coexpression relation, which would not be a thoroughgoing parallelism (which is highly implausible) but a selective one.

It also seems to me that panpsychism actually mitigates some of the problems associated with epiphenomenalism. The classic statements of this view make consciousness a radical emergent. By contrast, panpsychism holds that consciousness is a universal expression of its associated physical state. One might say that its role is just to testify to the pattern of physical events that occasion its fusions and its elementary manifestations. On this view, we should not expect consciousness to have a causal role in the world but simply to reflect patterns of that world.

Constitutive, not Causal. The foregoing are general arguments. Chalmers also makes some remarks specifically directed at accounts similar to combinatorial infusion-based DEP, of which I will only consider here what I take to be the most significant. Chalmers complains that since combinatorial infusion is a diachronic relation it is hard to see how it could be constitutive. As he notes, such diachronic relations are generally regarded as contingent and causal. Here, the defender of infusion must plead guilty, with an explanation. The explanation is that the world just is causally structured so as to support combinatorial mental infusion. We already know that the physical world is set up to enable combinatorial physical infusion¹⁰. But this is a feature, not a bug. It is what you get when you take seriously Mill's and Nagel's idea of 'mental chemistry.'

Although none of these replies is remotely definitive, it seems to me that they show that if panpsychism is to be taken seriously at all, then the abundant advantages of a combinatorial infusion-based DEP make it rather attractive and eminently worth further investigation.

Notes

- 1. The core arguments find canonical representations in Nagel (1974/1979); Jackson (1982); Kripke (1980, lecture 3); Chalmers (1996, especially chs. 3–4).
- A particularly fascinating example is discussed in Goodale and Milner (2004). The unfortunate subject, who suffered carbon monoxide-induced brain damage, is able to perform a number of complex perceptual tasks without awareness. For a detailed investigation of the relevance of various blindsight thought experiments to the problem of consciousness, see Siewert (1998).
- 3. A vaguely analogous situation might be the rich set of concepts developed by Christian thinkers in the late Middle Ages to describe the occult world of demons, angels, and witches. The familiar scientific example of phlogiston can serve as another illustration. A reasonably complex and empirically successful theory was developed around the notion of this nonexistent substance.
- 4. Others have also disputed the coherence of radical emergence, notably Galen Strawson (2006). Ithink, however, it is very hard to show that radical emergence is self-contradictory. A defense of at least the coherence of radical emergence can be found in McLaughlin (1992); I discuss Strawson's argument in Seager (2012a).
- 5. A full panpsychism requires that the mental be ubiquitous and fundamental. One might argue, as did Nagel, from the claim that conscious beings can be constituted out of any sort of physical constituents (e.g., antimatter vs. matter) to the claim of ubiquity.
- 6. In Mill (1843/1963, ch. 4) one finds a seminal discussion of radical emergence expressed in terms of the distinction between 'homopathic' and 'heteropathic' laws. These are closely analogous to Broad's intra-versus transordinal laws of nature.
- 7. See Humphreys (1997) for a general account of 'fusion emergence'; for a critique of Humphreys's approach see Wong (2006).
- 8. Does this claim violate Chalmers's principle of organizational invariance (Chalmers 1996, ch. 7)? It might or it might not, depending on how the fundamental mental features are distributed over the fundamental physical entities whose interactions and relational structures realize functionally definable systems. It might, for example, be the case that any system functionally equivalent to the human brain would occasion mental fusion. This mental fusion might or might not be phenomenally indistinguishable from brain-based fusion, or it might be in various ways and to greater or lesser extent phenomenally different (inverted spectrum cases fall under this possibility). Or it could even be that some systems functionally isomorphic to the conscious brain do not occasion fusion at all, leading to the possibility of zombies.
- 9. Prinz is a physicalist who endorses the strategy, discussed above, of defusing the problem of consciousness by showing how the explanatory gap is a cognitive illusion. However, his AIR theory is a perfectly intelligible account of the physical ground of consciousness and the phenomenal character of consciousness in its human form even if the physicalism is rejected.
- 10. Chalmers claims that classical physics does not support combinatorial infusion whereas quantum physics does. I think the example of the black hole, which is a classical phenomenon, shows that this is not quite right.

Bibliography

Broad, Charlie Dunbar. Mind and Its Place in Nature. London: Routledge and Kegan Paul, 1925. Chalmers, David. The Conscious Mind: In Search of a Fundamental Theory. Oxford: Oxford University Press, 1996.

Chalmers, David. "The Combination Problem for Panpsychism." *Panpsychism.* Ed. Godehard Brüntrup and Ludwig Jaskolla. Oxford: Oxford University Press, this volume. Dennett, Daniel. Consciousness Explained. Boston: Little, Brown, 1991.

Goodale, Melvyn, and David Milner. Sight Unseen. Oxford: Oxford University Press, 2004.

- Hameroff, Stuart, and Roger Penrose. "Conscious Events as Orchestrated Space-time Selections." *Journal of Consciousness Studies*, 3 (1996): 36-53.
- Hameroff, Stuart, and Jonathan Powell. "The Conscious Connection: A Psycho-physical Bridge Between Brain and Pan-Experiential Quantum Geometry." *Mind That Abides: Panpsychism in the New Millennium*. Ed. David Skrbina. Amsterdam: Benjamins, 2009. 109–127.
- Humphreys, Paul. "How Properties Emerge." Philosophy of Science 64.1 (1997): 1–17.
- Jackson, Frank. "Epiphenomenal Qualia." Philosophical Quarterly 32 (1982): 127-36.
- James, William. The Principles of Psychology. Vol. 1. New York: Henry Holt, 1890.
- James, William. A Pluralistic Universe: Hibbert Lectures to Manchester College on the Present Situation in Philosophy. London: Longmans, Green, 1909.
- Kripke, Saul. Naming and Necessity. Cambridge: Cambridge University Press, 1980.
- Levine, Joseph. "Materialism and Qualia: The Explanatory Gap." *Pacific Philosophical Quarterly* 64 (1983): 354–61.
- Loar, Brian. "Phenomenal States." Philosophical Perspectives, 4 (1990): 81–108.
- Lockwood, Michael. Mind, Brain and the Quantum. Oxford: Blackwell, 1989.
- McGinn, Colin. *The Mysterious Flame: Conscious Minds in a Material World*. New York: Basic Books, 1999.
- McLaughlin, Brian. "The Rise and Fall of British Emergentism." *Emergence or Reduction*. Ed. Ansgar Beckermann, Hans Flohr, and Jaegwon Kim. Berlin: De Gruyter, 1992, 49–93.
- Mill, John Stuart. "A System of Logic." *The Collected Works of John Stuart Mill*. Vols. 7–8. Ed. John M. Robson. Toronto: University of Toronto Press, 1843/1963.
- Mill, John Stuart. "Notes to James Mill's 'Analysis of the Phenomena of the Human Mind.'" *The Collected Works of John Stuart Mill*. Vol. 31. Ed. John M. Robson. Toronto: University of Toronto Press, 1869/1989.
- Nagel, Thomas (1974). "What Is It Like to be a Bat?" *Philosophical Review*, 83.4 (1974): 435–50. Rpt. in *Mortal Questions*. Ed. Thomas Nagel. Cambridge: Cambridge University Press, 1979, 165–80.
- Nagel, Thomas. "Panpsychism." Mortal Questions. Ed. Thomas Nagel. Cambridge: Cambridge University Press, 1979. 181–95. Rpt. in Panpsychism: Past and Recent Selected Readings. Ed. David S. Clarke. Albany: SUNY Press, 2004. 137–52. 137–52.
- Papineau, David. "Phenomenal and Perceptual Concepts." *Phenomenal Concepts and Phenomenal Knowledge*. Ed. T. Alter and S. Walter. Oxford: Oxford University Press, 2006.
- Prinz, Jesse. The Conscious Brain: How Attention Engenders Experience. Oxford: Oxford University Press, 2012.
- Rey, Georges. "Sensational Sentences." *Consciousness*. Ed. Martin Davies and Glyn Humphreys. Oxford: Blackwell, 1993. 240–57.
- Seager, William. "Emergentist Panpsychism." Journal of Consciousness Studies 19.9–10 (2012a): 19–39.
- Seager, William. Natural Fabrications: Science, Emergence and Consciousness. Berlin: Springer Verlag (Frontiers Series), 2012b.
- Searle, John. "Consciousness and the Philosophers." New York Review of Books 44.4 (1997): 43-50.
- Siewert, Charles. The Significance of Consciousness. Princeton: Princeton University Press, 1998.
- Stapp, Henry. "A Quantum Theory of the Mind-Brain Interface." Mind, Matter and Quantum Mechanics. Ed. Henry Stapp. Berlin: Springer, 1993. 145–72.
- Strawson, Galen (2006). "Realistic Monism: Why Physicalism Entails Panpsychism." Journal of Consciousness Studies 13.10–11 (2006): 3–31.

White, Stephen. "Curse of the Qualia." Synthese 68.2 (1986): 333-68.

- White, Stephen. "The Property Dualism Argument." *The Waning of Materialism*. Ed. Robert C. Koons and George Bealer. Oxford: Oxford University Press, 2010. 89–114.
- Whitehead, Alfred North. Process and Reality. New York: Free Press, 1929/1969.
- Wong, Hong Yu. "Emergents from Fusion." Philosophy of Science 73 (2006): 345-67.

10

Panpsychism and Neutral Monism

How to Make up One's Mind

SAM COLEMAN

Chalmers has helpfully distinguished a slew of combination problems.¹ His overall sense is apparently that *constitutive Russellian panpsychism* and *constitutive Russellian panprotopsychism* are broadly equally afflicted.² I aim to show that panprotopsychism is actually in much better shape than panpsychism, once one takes their respective combination problems into account. Panpsychism's distinctive combination problem, concerning the combination of subjects, reveals the theory as deeply unsatisfactory. The view I endorse— a form of panprotopsychism labeled 'panqualityism' by Chalmers—doesn't face the subject combination problem. On panqualityism the world is ultimately constituted of quality-instances, where we can usefully think of these as *unexperienced qualia*—properties just like the qualia we experience, only without anyone experiencing them.³ But since panqualityism does without the panpsychist's microsubjects, it must generate macrosubjectivity from scratch—this opens a new arena of combination problems specific to forms of panprotopsychism.

Chalmers believes panqualityism cannot provide the required reductive explanation of subjectivity, because it is vulnerable to a kind of conceivability argument. I'll argue (section 4) that panqualityism is not vulnerable in the way Chalmers suggests. First I'll explain what's wrong with panpsychism (section 2), and in between (section 3) I'll offer suggestions as to how panqualityists might deal with some of the other combination problems Chalmers mentions, as well as some he omits. The overarching thesis is not just that panpsychism is effectively a nonstarter, but that panqualityism has the resources to deal with its combination problems. In view of the advantages panpsychism and panqualityism share with respect to mainstream physicalism,⁴ this installs panqualityism as our best prospect for a theory of how to make up the mind.⁵ I start with some discussion and categorization of the combination problems.

10.1

i. The most important combination problems Chalmers exhibits include *the subject problem*, of how a number of subjects could synchronically constitute another subject; *the palette problem*, of how a handful of basic qualities could generate the vast array of macroqualities; *the quality problem*, of how qualities constitute other qualities at all; *the structural mismatch problem*, of how the micromental could constitute the structure of the macromental given that the micromental is isomorphic with the microphysical; *the unity problem*, of how disunified instances of micromentality come together to yield the unity we find at the macromental level; *the grain problem*, of how we get a homogenous phenomenal field at the macrolevel from discontinuous instances of micromentality; *the boundary problem*, of how micromentality is corralled into bounded units of consciousness.⁶ Some of these problems have variants applying specifically to panpsychism or panprotopsychism: I'll detail these as we go.

ii. We can distinguish two kinds of combination problem. A completed naturalistic account of mentality would mesh our best theory of minds with our best scientific theory, in particular with physics and neurobiology. Much of what we do as philosophers of mind is somewhat insulated from detailed scientific concerns, however. We do often employ scientific concepts as starting point, and we certainly *hope* that what we're devising isn't obviously inconsistent with accepted science. But often enough we're busy working on things from within the mind side—developing theories aimed at explaining aspects of mentality, and which are framed largely in mental or cognitive terms.

Consider as an example Rosenthal's *higher-order thought theory of consciousness*,⁷ which analyzes a conscious state as one that is the object of the right kind of occurrent thought. Naturally, Rosenthal would be dismayed if his theory turned out to be incompatible with our best neuroscience—were there, say, no feasible neurological candidate for the HOT monitoring mechanism. Yet it's clear that, broadly speaking, he *first* formulated his theory in mental/cognitive terms, and (for various reasons) only down the line is there any chance of seeing how HOT theory meshes with the science.

Meanwhile Rosenthal is bombarded with objections from philosophers. These are almost exclusively in mental/cognitive terms: It's alleged that

Rosenthal's theory doesn't capture the *explanandum*, phenomenal consciousness, because an unconscious thought plus an unconscious sensory state do not a conscious state make. Worse, it's claimed that Rosenthal's theory is internally inconsistent.⁸ Others say it cannot cope with possible content mismatch between HOT and sensory state.

These objections to HOT theory all take place within the arena of mindtheory, the immediate area wherein HOT theory aims to forge a coherent and illuminating position. They test its consistency or its fit with other things we believe about the mind, on that same theoretical level. Quite another sort of objection might allege that HOT theory cannot be neurologically implemented, or is in some other manner incompatible with established science.⁹ Objections to a theory of mind that flow from its fit with our best science I call '*bridging problems*'—they are difficulties around our building a bridge between a given theory of mentality and our scientific theories of the brain and physical world. The former sort of objection are 'internal problems'—they are difficulties alleged to afflict a theory of mind taken on its own terms, or within the field of theories of mind.

iii. It becomes apparent that some of Chalmers's combination problems are internal problems while some are bridging problems. The subject problem is internal—it concerns our concept of a subject and whether subjects could constitute another subject. The palette problem as Chalmers frames it is a bridging problem. Chalmers says the difficulty concerns a small set of qualities generating a 'vast array' of macroscopic qualities. That phrase suggests a problem of quantity: how do you get many (type distinct) macroqualities out of a few microqualities? The reason there are only a few microqualities is that Chalmers makes the microqualities isomorphic with microphysical properties, of which there are apparently only a handful. So the problem is: we want to make microqualities isomorphic with microphysical properties;¹⁰ that means only a few microqualities, so how do they generate masses of macroqualities? With the tie to physics removed this problem would be considerably less impressive: without the limited repertoire of microphysical properties we'd have no reason not to indulge in masses of microqualities. And if we have masses of microqualities, it won't seem so problematic to derive a vast array of macroqualities from them, taken just in numerical terms. There are enough ultimates—we could even have every macroscopic quale-type mapped to a token ultimate, or something like that. So Chalmers's palette problem is a bridging problem.

Chalmers distinguishes a further quality-related problem, which we might call the *production problem*. The difficulty is that we've no model for how qualities combine when these are separately instantiated, say by two distinct ultimates. If we have red and white in the same spot we can understand that as pink, Chalmers reckons. But if the red and the white belong to different items, how do we then get the pinkness? This difficulty appears closely related to another quality problem Chalmers doesn't directly consider, but which I find in Lucretius. It has to do with whether we *can* understand red and white in the same place as pink, or at least, with how exactly we do this. Lucretius compares combining ultimates of different qualities to manufacturing a perfume, saying:

> Among the first things that you need to seek Is an oil that is, so far as you may find one Odourless and emits no breath of anything. For this will least with harsh taint of its own Corrupt the scents concocted with its substance. For the same reason atoms must not bring An odour of their own in making things.¹¹

If you assign an ultimate a determinate quality, then as long as it continues to exist in the wholes it composes, which is a condition of its *constituting* those wholes, its quality must show up in—contaminate—them. So even if we can get past the production problem, and understand how separately instantiated qualities could interact to combine, we have an arguably more basic problem, of understanding how qualities that can combine actually do so. What does it even mean for two qualities to constitute a quality? Pink isn't red, and it isn't white either. One might expect that for red and white to survive in combination we would get as product a patch of red alongside a patch of white. Perhaps at a distance we might see that as pink, but it isn't, by itself, pink. We need some conception of qualities interpenetrating and yielding a new product, whilst nonetheless (somehow) persisting, corresponding to the combining of ultimates and their properties that also survive in the combination.¹² These latter two quality problems, then, concerning production and contamination, are both on the internal side—pertaining to the mechanics and dynamics of qualities taken by themselves, nothing really to do with science.

There may be one further quality problem, which seems to exacerbate the contamination problem. This concerns qualitative *incommensurability*. If ultimates have fixed qualities, just what set of microqualities *is* it that can be rearranged now as the smell of roses, now as an orgasm, now as a percept of the blue sky? These macroqualities seem so qualitatively *different*, it's hard to imagine generating them from some stable basic palette.¹³ This problem is most vivid when we limit the micropalette to a few qualities in order to fit with physics. But it isn't essentially a bridging problem; it would be

problematic even if we decoupled from Russellianism and went in for masses of microqualities.¹⁴

The *contamination problem* would be a challenge even if we didn't face apparently incommensurable macroqualities—qualities that seemingly couldn't come from the same ingredients merely rearranged. White and red don't seem *worlds apart* qualitatively, in the way that both do when contrasted with the smell of roses, and in the way these three qualia do, in turn, when compared with the feeling of a pinch on the forearm. Yet still we have work to do to grasp how red and white exist in—contaminate—their pink product. The *incommensurability problem* can be seen as making the *contamination problem* all the harder, or as an especially tricky aspect of it.¹⁵

What we require overall, when it comes to our micropalette, is a story which explains the coming together of separately instanced qualities, explicates the very notion of qualities constituting other qualities, and defuses the sense that certain macroqualities are so qualitatively different from one another that they couldn't derive from a stable set of basic ingredients. If we can do these things we will presumably also answer Chalmers's question about how the vast array of macroqualities is produced.¹⁶

iv. *Structural mismatch* is a bridging problem. If microexperiential structure matches microphysical structure, then it seems macroexperiential structure is restricted to isomorphism with macrophysical structure: yet macrophysical and macroexperiential structures differ, Chalmers suggests. Were it not for the matching of microexperiential structure to microphysical structure, we would presumably be free to envisage microexperiential structure, so this problem concerns meshing our account of the mind with science.¹⁷ Chalmers's *grain problem* is something like an internal analogue of the *structural mismatch problem*. It has little to do with science, stemming only from the thought that microexperiential instances are *discontinuous*, while macroexperience, supposedly constituted by them, is continuous.¹⁸

v. Finally, the *unity problem*, and nearby *boundary problem*, are internal. The question is as follows: If you have a phenomenally unified and bounded experiential field, how is that constituted by discrete instances of micromentality? Since for panpsychists microqualities are had by microsubjects, this difficulty is for them entwined with the *subject problem*: how are separate microsubjects to constitute a macrosubject with its own, single, experiential field?

vi. There's a case for considering internal problems as more pressing than bridging problems. It seems that if we're unable to put a coherent theory of mind on the table in the first place, taken on its own terms, then we needn't venture to check how it matches with the science. The theory is already hopeless. Of course lack of mesh with science is also serious. But since there is usually more than a single option for understanding the science of the moment, and the dominant scientific account in an area at a time is highly liable to shift, this makes lack of bridging arguably less urgent than an internal clash for a theory of mind; at least, as regards widely-accepted and stable posits of mindtheory, like phenomenal consciousness, or the existence and nature of macrosubjects, our prime concerns. Given lack of bridging, one could at least hope for a change on the scientific side to remove the obstacle. In any case, even if a mind theory *fatally* fails to bridge, running against some scientific bedrock, it seems that to get that far along the proving process it had already to be in decent shape internally. Thus internal problems have a certain theoretical priority. The really big immediate questions for panpsychism and panqualityism, accordingly, concern their most serious respective internal problems: for panpsychism, whether it can deal with the subject combination problem; for panqualityism, whether it can generate subjects from nonsubjects. These topics form the backbone of our discussion.

10.2

i. Examination of its distinctive internal problem, the subject combination problem, will lead us to seriously question the basic theoretical motivation for panpsychism.¹⁹

ii. James is widely cited on the subject problem,²⁰ but Lucretius had his eye on this one too. Were ultimates subjects of experience, he argues:

their unions and combinations, Would make nothing more than a crowd of living things Any more than men and cattle and wild beasts By combination could make anything.²¹

Taking James and Lucretius together, we can discern a positive and a negative subject problem for panpsychism, which correspond, roughly and respectively, to Chalmers's *unity problem* for panpsychism and what he calls the *subject-summing problem*. The negative problem, subject-summing, is effectively an *explanatory gap*: no amount of talk of subjects coming together seems to entail anything about any further subject. So it doesn't appear that panpsychism can account for the constitution of a macrosubject, which was certainly the aim of constitutive Russellian panpsychism.²² The positive difficulty is something

like a genuine metaphysical stumbling block or apparent *impossibility*: How *could* you hope to produce a phenomenally unified, single-perspective, subject by assembling a group of subjects each of which essentially has its own perspective? It really doesn't seem that you could, in a constitutive manner. Constitution requires the relationship between parts and whole to be synchronous, and means that all there is to the constituted phenomenon are the entities said to constitute it, their properties, and their relations. That entails, for panpsychism, that at a time the existence of a single macrolevel perspective—a conscious point of view like one of ours—*is nothing but* the existence of a group of (micro)subjects, each with its own perspective. But a group of subjects looks like a crowd, and a crowd is not a unified conscious mind. I've developed the positive problem elsewhere,²³ but as the explanatory gap problem for panpsychism is more widely known, and suffices for our purposes, I'll concentrate on it here.

As currently elaborated, the subjects-summing problem gets embedded in the following argument:

- (1) If panpsychism is true, the existence of a number of microsubjects with certain experiences necessitates the existence of a distinct macrosubject.
- (2) It is never the case that the existence of a number of subjects with certain experiences necessitates the existence of a distinct subject.
- (3) Panpsychism is false.²⁴

The support for premise two (the subjects-summing problem) is the alleged fact that:

(*) For any group of subjects (with certain experiences), it is conceivable that those subjects exist (with their experiences) and no other subjects exist.

I find this argument fairly powerful. But I don't think it quite gets to the bottom of the deep problem panpsychism faces in this vicinity. This deep problem flows from the following principle:

(**) Fundamental intrinsic properties help to explain their macroscopic instances.

This principle is metaphysical on its face, but it also has a methodological aspect. Fundamental posits, especially of intrinsic properties, must earn their explanatory keep. Specifically, we posit a fundamental property for the purpose of accounting for its higher-level instances.²⁵ Mass, charge, and extension all do this job, indeed it explains their being attributed at the basic level at all.

We don't make otiose fundamental posits, so any posit that doesn't earn its explanatory keep should be discarded. More properly, it should not have been entertained in the first place. This is the situation we face regarding the panpsychist's posit that ultimates possess the intrinsic property of subjectivity.

iii. The reason the subjects-summing argument doesn't quite touch the depths of this problem has partly to do with the fact that it talks not in terms of *explanation*, but of *necessitation*. This permits the panpsychist a certain kind of sidestep. If no mere assembly of subjects necessitates a further subject, but, as prospective panpsychists, or theorists rightly bent on giving the model a fair chance, we aspire to such necessitation, a natural suggestion arises, in the form of Goff's *phenomenal bonding*, tentatively endorsed by Chalmers.

Goff's idea, I believe, is that we've perhaps focused too exclusively on microsubjects and their properties, without thinking imaginatively enough about what relations among them might accomplish.²⁶ Perhaps the addition of some special relation to a set of subjects might supply the necessitation of a macrosubject. So Goff's proposal is that a macrosubject forms when a set of microsubjects is said to 'phenomenally bond': a subjective analogue, it seems, for the bonding among atoms which forms molecules.

But, we may reasonably inquire, concerning a set of subjects from whom a macrosubject is thus produced, what exactly *is* their phenomenally bonding, just what does it amount to? As far as I see, each microsubject contributes some *experiential contents* that then get enjoyed by the macrosubject. This much makes sense: it's not obviously incoherent that a third person could have experiences now, constituted of some of what you and I are experiencing; why shouldn't we contribute experiential contents to this individual? Maybe we can even keep on experiencing our separate contents meanwhile. This all seems (at least) intelligible under the heading 'telepathy.'

The difficulty for phenomenal bonding comes not on the experiential content side, but on the subject side. We need not only to provide the new macrosubject with contents to experience, we are required to *manufacture* this macrosubject in the first place: the point of view for whom the contents are to be like something. And there's nothing in the sheer idea of phenomenal bonding that tells us how discrete subjects produce a new subject. What 'groups' those subjects' experiences together in a new phenomenally unified perspective, a new bounded field? Certainly nothing about having microsubjects already in play explains this. Each of the experiential packets—the contents—to be proffered to the novel subject belongs already to a point of view, and we are imagining those to remain intact on this *constitutive* model. Perhaps each content packet extended to the new subject carries with it (somehow) a 'quantum' of the subjectivity of its previous owner.²⁷ But then, clearly, we would simply end up with multiple perspectives bunched together in the new spot (the construction site of the prospective macrosubject), one for each packet of experiential content proffered by a microsubject: for, surely, any 'quantum' of subjectivity implies a subject, so implies a perspective. We would then need, anew, to explain how these several perspectives ('subjective quanta') added up to a single unified one. The subject problem thereby respawns, which is the first sign of a nasty regress: nasty because we've made no explanatory progress.

iv. Chalmers attempts to put flesh on Goff's idea:

A natural candidate here [for the phenomenal bonding relation] is the co-consciousness relation: a relation such that whenever it relates two phenomenal states, they are experienced jointly. When this relation holds among the states of distinct microsubjects, those states will be experienced jointly by a new subject.²⁸

But this is to *describe* the (desired) outcome of a certain process, without telling us at all *how* it is meant to be achieved. Co-consciousness requires a subject: it's consciousness *for a subject* of some items. That makes *being co-conscious* relevantly like *being co-punched*, in that when two things are co-punched, we must ask: *By whom*? When we drag two experiential packets out of respective microsubjects, whence does the new subject come for whom they are to be co-conscious? To say that experiential packets are related now by co-consciousness is certainly to *imply* that a new subject has come into being for whom they are phenomenally unified, but it is not to tell us how this happens, nor whether it is possible—the things we wanted to know. In the absence of further positive content, what this notion of phenomenal bonding really amounts to is a *schema*: it specifies what an explanation. It is a mere black box.

v. What if we try just to take phenomenal bonding at face value? We can readily enough imagine that when a number of microsubjects get into the requisite relationship—whatever it is—a macrosubject pops out of thin air. Thus phenomenal bonding, somewhat by stipulation, might supply the *necessitation* of a macrosubject: what was at issue in the subjects-summing argument. With phenomenal bonding added to the account, and so to the background of one's conceiving, one might then no longer be able to conceive of a set of subjects getting into the prescribed relationship without a macrosubject forming. But the glaring truth about this 'explanation' is that the fact of the phenomenally bonded ingredients' being *subjects* plays no role whatever. All the work is done by the phenomenal bonding relation: it is a relation such that, by definition, it generates a macrosubject. It seems that we could as well imagine panqualityist-style *subjectless* qualitative patches as our ingredients, and posit a bonding relation such that they not only pooled sensory contents, but generated a subject to experience these. We could then talk, with Chalmers, of "a relation such that whenever it relates two [panqualityist] states, they are experienced jointly . . . by a new subject."²⁹ It sounds just the same: there is as much—as little—explanation of the constitution of the novel subject on both accounts.

So even if phenomenal bonding could fill the lacuna around necessitation that looms large in the subjects-summing argument, it does nothing as yet to tell us what the explanatory role of the panpsychist's microsubjects is. Thus far, the intended role of subjects in the constitution of a subject is entirely opaque. That is the really deep problem for panpsychism: What is the principled motivation for positing microsubjects in the first place, just what explanatory work do they do? The fact that we are tempted to appeal to phenomenal bonding shows that panpsychists lack an answer to this question.

vi. What then would a good panpsychist explanation of 'phenomenal bonding' look like? It's informative to consider the case of atoms bonding into molecules. Relations do a lot of work there; for instance with the oxygen's sharing of the hydrogens' electrons in the formation of water. But the important point about such relations is that they visibly *flow from* the intrinsic natures of the *relata*. This is generally the case with relations, in fact.³⁰ It is because of the relative looseness of hydrogen electrons, coupled with the convenient gap in the oxygen's outer shell, that electron sharing happens so readily in the constitution of water. For phenomenal bonding to work, we would need some analogue of this sort of explanation. It would be taking into consideration the intrinsic features of microsubjects that suggested to us the mechanism for their phenomenally bonding—a subjective equivalent of electron sharing. The case with subjects is in reality exactly the reverse: it is because panpsychists *cannot* see how subjects could come together, given their intrinsic properties, that the supplement of phenomenal bonding is broached. We have here a relation devised precisely to remedy the obvious defects of its putative relata. This is bad news for panpsychism: it strongly suggests that microsubjectivity is (at best) explanatorily irrelevant to the constitution of macrosubjects, which in turn rules it out as a fundamental posit, according to (**).

vii. My diagnosis of panpsychism is as follows. It becomes tempting thanks to the admitted starting power of the intuition that *from ingredients lacking entirely in subjectivity, we could not a subject produce*. However, panpsychists are guilty of sliding from this thought, to the doctrine that *if we just add subjectivity to our microingredients, all will be well*. Yet this transition is clearly unjustified, as it stands. For it may be that, even were the starting thought correct, it would not help to add subjectivity to our base. That is in effect what is demonstrated by the dialectic above, and the resort to phenomenal bonding: the posit of fundamental subjectivity as an intrinsic property of ultimates has yielded no progress in our attempt to account for macrosubjects.

The panpsychist's starting intuition, concerning the impossibility of deriving subjectivity from the nonsubjective, runs up against another intuitive principle, which seems at least as powerful: that subjectivity, far from being a diffused sort of 'stuff' or generalized property, inevitably comes in the form of certain quanta, namely *subjects themselves*. The self-contained nature of these quanta, their phenomenal unity and boundedness, makes them singularly unsuited to the constitution of any further entity.³¹ So while panpsychists suffer the intuition that subjectivity *must be there* in the microbasics, the form in which it unavoidably occurs, packaged as subjects, at once blocks the hopedfor explanatory payoff. Panpsychism is thus crushed between two irreconcilable intuitions.

It seems these two intuitions cannot both be correct. Yet while it may be hard to envisage a subject forming from nonsubjects, the idea that subjectivity could exist somehow *in general*, as a blanket quantity—apt to being broken down and reformed, like dough—seems more obviously incoherent. At most we picture a universe-subject when we think along these lines; but such an entity still has undeniably a point of view, and, as a corollary, creates a difficulty for the manufacture of subjects of our level. As long as the metaphysical solidity of subjects is acknowledged—a driving factor behind panpsychism, of course—panpsychism cannot, it seems, succeed. Panpsychists would have to embrace emergentism, owning that, with the assembly and phenomenal bonding of microsubjects, 'it just happens' that a macrosubject forms. Of theorists who take this route, we may even more properly ask why they require the posit of fundamental subjectivity.

viii. Since microsubjectivity is at best an idle wheel in the explanation of macrosubjects, even on the most promising panpsychist account, it is not a posit we should go in for. This means a rejection of panpsychism. We should retain qualities in micro-ontology, but deny that they require subjects to experience them. *Panqualityism* has thus a more pared-down ontology, and the question is whether it is adequate to the phenomena. This theory faces still the other combination problems, as well as a new set of its own: it must generate subjects and awareness from ingredients lacking both properties. This appears, prima facie, a pair of new internal problems at least as formidable as the subject problem was for panpsychism.

10.3

i. Light of the subject problem, we've the following combination problems still in play: *the palette problem; the structural mismatch problem; the unity problem; the boundary problem; the grain problem; the production problem; the contamination problem;* and *the incommensurability problem*. Additionally, we have a new *awareness problem,* which panqualityism faces due to stipulating that microquality instances are not conscious.³²

Isn't there also a new subject-related problem, of how subjects are constructed from nonsubjects, for panqualityism?³³ I don't think so, actually. For our purposes we may take a subject minimally to be an *aware entity*—anything that has conscious awareness. It follows that showing how there can be subjects, in the minimal sense, reduces to the problem of accounting for awareness.

The unity problem now splits in two: the *qualitative unity problem* is the same as, or will be solved in the same way as, the production problem. If we understood how quality instances tokened by distinct particulars could form into a combined single quality instantiation, it seems we would have given an account of how qualitative unities form. The unity problem for panpsychism concerned unity when there are multiple ultimate-*subjects* in play, with discrete unified fields of experience. We no longer have those subjects in play, only instances of quality belonging to different ultimates. There *is* a story to tell about how qualities, but put this way the difficulty seems equivalent to the production problem. So qualitative unity and production problems appear close enough to be counted as one.

The sense of 'unity' operative in the unity problem for panpsychism is *phenomenal unity*, the unity of the experiential contents *given* to a single subject. All the elements I experience can be grouped together phenomenally in what we can call my 'overall experience.' They are also separated phenomenally from the elements you experience. The challenge of accounting for these features hasn't gone away. When we come to the panqualityist mechanism for awareness, we'll see that it can naturally be used to account for the *phenomenal unity* and *boundedness* of macrosubjective experience.

ii. The production/qualitative unity problem: How do separately instantiated qualities (e.g. the redness and whiteness of distinct particulars) yield a combined quality? If there is a genuine problem here, then we perhaps need to reject the premise that the qualities are separately instantiated, or remain separately instantiated. Chalmers sees no problem with co-instantiated qualities combining. So if separate instantiations can become nonseparate in the combination of the qualities, we may have a solution. Chalmers discusses a 'quantum holist' picture where two or more ultimates can get entangled, gaining properties that then entitle us to treat them more or less as one item. Now suppose that two ultimates presently unentangled instantiate red and white, respectively. Then it might be reasonable to consider them, once entangled, as instantiating pink: the ultimates entangle, and so do their qualities. You can no longer take a red and a white 'reading' separately, as it were; rather the two qualities now have to be treated as a block. Some even interpret entanglement as the two (or more) entangled things literally becoming one thing.³⁴ If that's plausible, then the formerly separately instantiated qualities-redness and whiteness-are now co-instantiated, and the puzzle appears solved.35

Chalmers raises two major concerns about quantum holism, for panpsychists.³⁶ First, entanglement might be taken to unite the whole universeespecially since it likely emerged from a singularity.³⁷ This, in the context of panpsychism, would lead to a single universe-sized subject, and the *decomposi*tion problem: how do you get from that universe-subject down to us?³⁸ Second, Chalmers writes, "The structure of the quantum state of brain-level systems is quite different from the structure of our experience."39 This seems to be a bridging problem, a quantum structural mismatch problem. I have specific things to say about structural mismatch later. But for now we can observe that these two problems are significantly diminished without panpsychism. Panqualityism plus quantum holism won't imply a universe-subject, because ultimates aren't subjects, on this view, so their entanglement doesn't imply ever-bigger subjects. What entanglement might imply instead, under panqualityism, is a universe-wide entangled web of qualities. The universe could be conceived of as an enormous enqualitied fabric, with each quality instance being deeply related and entwined with all the others.⁴⁰ The quality of a given coordinate in that web to some extent supervenes on what qualities are present at all the other coordinates, so yielding a massively holistic qualitative universe.

We can achieve this result thanks to a difference in how the panqualityist is employing entanglement here, as compared with the quantum holist panpsychist. That panpsychist utilizes entanglement as the mechanism of subject combination. All we are doing is using it to overcome the obstacle of separate instantiation that Chalmers sees as blocking quality combination. When qualities are to combine their bearers plausibly have to be entangled, we may say. But we can frame this as a necessary but not sufficient condition of qualitative combination. This means that not every entanglement entails the production of a single unified quality. So if the whole universe is entangled, we needn't say that it instantiates but a single quality. Of course, we will need at some point to speculate as to the physical mechanism that, together with co-location, effects combination. But note that we have now moved from an internal problem how could separately instantiated qualities combine? (answer: they are not separately instantiated, in the relevant sense)—to a bridging problem: what is the physical analogue of quality combination? Since bridging problems are less urgent than internal ones, this represents a modicum of progress.

Because panqualityism, unlike this quantum holist panpsychism, doesn't envisage the structuring or entanglement of ultimates to constitute the structuring of consciousnesses into larger conscious wholes, the scope and structure of our experiential fields is an issue that floats to some degree free of questions about the more basic structure of the physical universe. The panqualityist has one job explaining how microqualities combine, and a separate task to account for consciousness of qualities, plus the unity and boundaries of macroconsciousness. This unburdens the panqualityist of the universesubject worry. But it also means there's no problem that the quantum structure of a brain doesn't match the structure of experience. By hypothesis, something extra is needed for experiences—aka awareness of the qualities—and the panqualityist could aim to manufacture an awareness relation that operates on entangled groups of ultimates at the macrolevel, and at the same time structures and defines the experiential field. This would amount to a panqualityist macroscopic 'cutter' of the brain-level or universe-level entangled quality-web, slicing in just the right places to yield fields of qualities corresponding in scope with those of our conscious acquaintance.

iii. Before investigating that cutter, I want to address the contamination problem. We can make sense of red and white sensory qualities adding up to pink, I suggest, in just the sort of intuitive way that we understand that if we mix a red and a white dab of *paint* we'll get a pink dab of paint in that spot. We're in the business of building a *constitutive* model of macroexperience, so we want to be able to say, of the pink quality, that its pinkness at a time is nothing but the redness and the whiteness, and their relationship, which requires in turn that these qualities survive in the whole. Clearly they don't survive as they were before, in separated form. Now that their particulars are entangled, the qualities are in a sense co-located. We can think of the qualities as surviving in their *contributions* to the pinkness, as in a sense the spin (etc.) of entangled particles survives from pre-entanglement as an aspect of the novel state.⁴¹ Analogously, distinct physical forces are present in complex real-life situations as the contribution each makes to the result. Intuitively, if you *were* able to remove the red quality from the blend, you would no longer have pinkness, and likewise for the white: so they are still present. Have the red and white survived with their qualitative identities intact? Yes: that's the only explanation of the ongoing fact that we have *this pink*. It takes precisely this red and this white to make this pink. But still, you can't find red on its own—or white for that matter—given their conjoint state.

If this is cogent, then we have an intuitive model for the constitution of qualities by other qualities,⁴² and a picture of how qualities 'contaminate' their wholes. They are still present, even though—in our example—as elements of a newly formed pinkness, red and white are now in a sense more dependent on the entangled whole. Notably, then, while we couldn't intelligibly make subjects constitute other subjects, we can intelligibly make qualities constitute other qualities. This represents panqualityism's essential advance over panpsychism.

iv. This line of thought raises the palette problem-how can we derive the masses of macroqualities just by blending a few microqualities (even if we understand the blending)? For a start we can note that this bridging problem takes a certain noncompulsory view of the scientific picture. It's true that if we concentrate on conventional microphysical quantities like mass, charge and spin, there don't look to be many slots available for qualities. Even if we ruled that it's particle types which each possess a distinctive kind of quality, that won't get us terribly many determinates. Given that we're likely to discover more particles, we can expect that number to increase somewhat. Another move is to consider matters in terms of string theory.⁴³ Strings can potentially vibrate in infinitely many ways, and perhaps each vibration corresponds to a different quality.⁴⁴ There'd be room then to include olfactory qualities, colors, and so on, all as having basic instantiations. There would still be much to do to work out how they combined, but the numerical challenge of Chalmers's palette problem wouldn't seem so pressing.⁴⁵ Anyway it's not clear how pressing that challenge is, even within the constraints of a few basic qualities. Given a few basic elements, these can clearly be combined in an enormous variety of ways: for any combination of one instance of each of the basic set, we can add a further instance of one of the basic members, presumably altering thereby the quality of the whole. In that way we can see there are potentially limitless places to go merely numerically.

v. This suggests that Chalmers's *palette problem* is in the end best understood as the *incommensurability problem*. Just what restricted set of microqualities is it that in recombination can yield now a pure blueness, now the smell of roses? We have perhaps gained some sense of how qualities can contaminate their products, but they must still contaminate them *as such*, as the qualities they are. The problem is that some macroqualia are apparently so *unlike* some other macroqualia that we can't imagine them having ingredients in common. This is the relevant sense of 'diversity': qualitative diversity.

I think the answer will require radical reconceptualization of our qualityspace: discarding the idea of discrete modalities, and coming to think of phenomenal qualities, of all kinds, as on a continuum, in the way we think of the colors. So just as it's possible to move across the color spectrum in tiny, almost undetectable steps, it must be possible to move from tastes to sounds, sounds to colors, and so on, via equally tiny steps. Tiptoeing between modalities already seems conceivable in certain cases, perhaps even actual. We know that what we experience as 'taste' is really some kind of fusion of qualia sourced from the nose and from the tongue (to separate these just eat something while pinching your nostrils). Perhaps we don't routinely notice this because of the qualitative overlap between olfactory and gustatory qualia. Again, sometimes a *thump*, especially experienced while falling asleep, is not clearly distinguishable as *felt* or as *heard* (it doesn't appear to be both)—overlap between tactile and auditory qualia seems intelligible. One thinks also, in this connection, of the experience of deep bass drumming. It even strikes me as plausible that tactile qualia are just (qualitatively) more 'forceful' or 'solid' counterparts of 'thinner' auditory qualia.

To address qualitative incommensurability we must stretch to conceiving of such continuities as the rule rather than the exception. Hartshorne ably defends this 'continuum hypothesis,' and for want of space I refer the reader to his discussion.⁴⁶ Let me only mention here his helpful idea that, where two sets of our qualia stubbornly appear absolutely different (as perhaps with taste and color qualia), this may be an artefact of a missing 'intermediary' modality that we lack (perhaps it is not evolutionarily useful for us to have it). So imagine that the qualia of the shark's electric sense are the missing modality in question: it would then be possible to stone-step from visual qualia to shark electric qualia through to taste qualia.⁴⁷

To further motivate the continuum hypothesis I offer the following small, hopefully suggestive, thought experiment: Imagine a creature whose qualiaspace featured only (what we would call) colors. Though possessing our five external senses, their qualitative products in its consciousness would just be colors of different sorts, with no color that features in 'audition' featuring also in 'vision,' and so on. We can get some grip on this being's mental life by thinking about the way that heat—thought of primarily as a tactile quality—can feature in visual experience as red or orange. Now we just have to imagine that the creature, in touching a warm surface, experiences these visual qualia *only*, instead of the tactile ones we feel. Plausibly this creature would come to conceive of the qualia corresponding to its various senses, what for us would just be different kinds of colors, as qualia belonging to irreducibly different 'modal spaces.' This would seem the likely result just as long as no particular color featured in more than one modality. We can imagine that the creature's tactile sensations are all varying shades of red, vision presents only blues, smell the greens, and so on, with the places where these qualities (for us) overlap conveniently screened out by the organism's evolution, to prevent confusion (just as we might hypothesize that we helpfully screen out certain overlaps between the qualia of our various 'modalities'). The distal stimuli, as well as the different transceivers by which we absorb signals from them, are all too apt to contaminate our conception of the qualia they elicit, generating misleading impressions of absolute difference. Perhaps a being with a qualia-space correspondingly greater than ours as ours is greater than the color-only creature, would conceive of human qualia as belonging to a single 'modality.'

If the continuum hypothesis is correct, then there isn't any genuine incommensurability between different kinds of qualities—differences are always of degree rather than of kind. We might well *think* there are incommensurables, because we lack some areas of quality-space that would join up the qualities in question. It is not inconceivable, then, that just as (I believe) we could entertain the idea of a basic set of colors that in recombination could get you to all corners of the color spectrum, there might be an 'intermodal' quality set that could take you to all corners of qualia-space as we know it, and beyond.⁴⁸ We will then understand qualitative identity and difference in terms of the numerical identity or difference of these underlying components. The qualities required to do this job are likely not directly conceivable for us, being as they must lie 'in between' all the qualia we know of (they are present in our qualia only as so many myriad trace contributions). But I have some hopes of a genius who will figure out how we may *qualitatively deduce* them, by triangulation from the qualities of our acquaintance.⁴⁹

vi. What of awareness? I favor a higher-order thought theory, where a HOT's suitably representing a sensory state constitutes that state's being conscious.⁵⁰ We might envisage a panqualityist world, a web of qualities, with the HOT systems in brains, by representing other bits of these same brains, enabling consciousness of certain tiny portions of the material universe. Rosenthal's notable insight regarding consciousness is that a conscious state is one the subject is aware of being in. This awareness is plausibly captured by the notion of mental representation of the conscious state, which swiftly leads to something like HOT theory.⁵¹ What the panqualityist incarnation has as advantage over conventionally physicalist HOT theory is the unreduced presence of

qualities—thus it has no need to account for the generation of qualities from the nonqualitative, nor to eliminate qualities (as is, arguably, the practice of conventional HOT theory and its kin). The HOT component of panqualityism is (almost⁵²) solely charged with producing *subjective awareness* of qualities. In section 4 I tackle the abiding sense that a HOT-based account is not up to even this task; for now we leave the issue of awareness behind.

vii. This *HOT panqualityism* has the resources to treat the unity and boundary problems. Now, if one's qualitative states were made conscious by a *set* of HOTs directed at different elements, there would, as Chalmers notes, ⁵³ occur a problem as to how these qualitative elements became phenomenally unified experienced together. It wouldn't follow from having this thought, and that one, about different qualities, that we had any thought about both qualities together. Yet we're conscious of all and only what HOTs target. So we'd get no unified qualitative consciousness as a result. The solution is to posit a single *very big* HOT for each of us, at a time: a complex thought taking in all the qualities we're synchronously aware of, perhaps a big conjunctive thought. Then we'll experience all the relevant qualities together, and we have unity.⁵⁴

We also have *boundaries*: your HOT covers a certain range of qualities in *your* body, and those are the ones of which you're conscious. Mine does likewise for a set in *my* body, which your HOT does not target: this is a simple matter of physiology. The relevant HO representation will likely require a nontrivial amount of neurological integration between representing and represented states.⁵⁵ This ensures I can no more HO represent your sensory states than I can digest the alcohol you consumed last night. These states of yours are simply out of my reach. Since we are each conscious of all and only that which our respective HOT systems target, we'll get two separate, bounded, unified fields of consciousness, on this model.

If we envisage a universe-wide web of qualities, really one structured field, then the HOT systems we bear are *cutters*, chopping, in each case, a defined patch out of the overall web and producing awareness of it; hence generating subjects—*loci* of awareness, each at the center of (i.e., phenomenally *given*) a bounded field of qualities. Panqualityism deals with the qualities of which we're aware; the HOT component, as well as providing awareness at all, fixes the experiential field and its properties, like boundedness. It's all too tempting to compare the HOT systems to *spotlights*, illuminating minute areas of the panqualityist universe.

viii. If the universe is a continuum at the microphysical level, this permits a considerable amount of graininess at the macrophysical level nonetheless. A neuron, for instance, will on the field conception show up as a node, or massive *knot*, in the universe's quality-fabric. Though *ultimately* continuous with its surround, such a node can be treated for certain purposes—like those of measuring its electrical potential—as an isolated unit. Neurological accounts are framed in terms of the commerce among such units (together forming *circuits* and systems). So even in the panqualityist universe there remains a job to square this macrophysical graininess with the smoothness of experienced qualities, especially given the apparent importance of neuronal level goings-on to consciousness.⁵⁶

Rosenthal suggested several years ago that HOT theory might assist with the grain problem: "The mental properties of our sensations appear ultimately homogeneous to us simply because the way we are conscious of them [i.e., the HO representation] smooths them out, so to speak, and elides the details of their particulate, bit-map nature."57 With a HOT and sensory state in play in a given case, the subject's awareness is effectively placed at a little distance from its object. This creates just the space we need for a small appearance/reality gap. Without going so far as to declare that the state one is conscious of in reality lacks qualities, ⁵⁸ we have nevertheless the room to say it perhaps lacks the (macrolevel) smoothness it appears in consciousness to possess. Thus the grain problem is finessed: we might claim that there is discontinuity at the macroexperiential level which does not show up in awareness thanks to the 'clumping' or smoothing effect of HOTs. This could work, in part, as follows. We are conscious of all and only that which our HOT targets. Thus if the HOT selectively targets discontinuous, even widely-distributed, brain features, we will be conscious of these without the gaps, since the gaps are by hypothesis not targeted, hence not represented. Being conscious of these items without the gaps between them is to be conscious of them in a continuous field. Analogously, widely spaced TV cameras, focusing on distinct parts of a scene, supply a spatially continuous image on the television screen. What is not dealt with by this explanation is the microphysical grain within (in the constitution of) a given macrophysical brain portion (e.g., neuronal activity node): Why are we not conscious of its microqualitative texture? Perhaps this is a matter of the relatively low 'resolution' of HOTs.59

So, although Lockwood avers that "no literal sense can be attached to the notion of the conscious mind being distanced, in this fashion, *from itself*,"⁶⁰ we can on HOT panqualityism in fact stand to the brain's structure (in consciousness) somewhat as one does in viewing a newspaper photograph:⁶¹ we perceive not the 'dots,' only the image.

ix. Structural mismatch seems susceptible of similar treatment. Lockwood may be right that "what is ostensibly lacking . . . is even the most approximate isomorphism between states of awareness and the underlying physiological

goings-on."62 But on HOT panqualityism, the requisite mirroring is not between what we are aware of and brain structure simpliciter, but between the HOT mechanism *along with* the sensory percept it carves up, and brain structure. Given a relatively 'raw' qualitative feed from the sensory systems, the HOT mechanism's job is to complete the preparation of a percept fully ready for presentation to the subject (in awareness), and fit for the task of negotiating the world. It's no surprise, given an environment of significant medium-scale dry goods (threats, food, etc.), that the HOT system should have evolved to 'gloss over' (in Lockwood's phrase) the complexity, and amend the structure, of the pre-conscious raw qualitative feed; all in the cause of helpful isomorphism between the conscious percept and the subject's environs, not her brain.⁶³ This corresponds, effectively, to Chalmers's suggestion that the sought-after mirroring between experiential field and brain structure is at the *informational* level. The proposal is in the spirit of Lockwood's observation that a functionalist account might have the wherewithal to abstract from the nitty-gritty of physical implementation when determining the structure of the conscious field. Lockwood dismisses the proposal because "Functionalism may have some plausibility in accounting for mental structure but, on the face of it, fails utterly to account for phenomenal content."64 Yet this objection doesn't touch two-pronged HOT panqualityism, the right wing of which supplies qualitative content in unreduced form. The functionalist (HOT) unit swoops in only to supply, and regiment, awareness of this qualitative feed. 65, 66, 67

x. One's natural next thought might be to wonder whether panpsychists could avail themselves of this useful HOT apparatus. Indeed they could. Panpsychists must anyway invoke some kind of special relation among sets of ultimates, superadded to their property of consciousness. For a panpsychist must explain an underappreciated problem—why subjects are bounded as they are: we do not presumably experience all the conscious ultimates within our bodies; and even if we did, panpsychists respect commonsense ontology enough to carve us (human subjects) off from one another experientially. The panpsychist must therefore posit some relation that all and only the conscious ultimates comprising *my* consciousness stand in, likewise for *you*, and so on.⁶⁸ What's to prevent panpsychists adopting the HOT mechanism in its capacity as 'cutter'? They might then also help themselves, it appears, to whatever power this supplement possesses when it comes to the grain and structural mismatch problems.

Yet this move promises more harm than good for panpsychists, for it serves only to underscore the essential idleness of the posit of fundamental subjectivity. Having ultimate-*subjects* in play is thus revealed as doing nothing to help with understanding *either* the constitution (see section 2) *or* (now) the structuring of a macrosubjectivity. Worse, in turning to the HOT mechanism to treat grain and structural mismatch, the panpsychist even outsources some of her account of the *phenomenal* state of macrosubjects. Theoretically more elegant panqualityism simply has the sort of relation a panpsychist might appeal to here do double duty for awareness. All that remains of panpsychism, by now, is the plaintive cry that *without microsubjectivity we could not generate macrosubjectivity*. But, as we saw, this thought is a dead end: the addition of microsubjectivity did not help in accounting for macrosubjects.

10.4

i. The foregoing discussion suggests the potential value of HOT representation, in combination with panqualityism, when it comes to treating some of the more intimidating combination problems.⁶⁹ But this promise is for naught if there's reason to think that awareness cannot be analyzed by higher-order thought. The slippage between sensory states and awareness provided by the HOT mechanism, and the structuring it offers of the contents of awareness, appear conditional on HOTs sufficing for awareness in the first place. But Chalmers avers that any such attempt to 'functionalize' awareness must fail, since it will face its own variety of zombie.

The relevant zombie argument starts from a panqualityist world description: a quality web, including (or implying) the functional structures in brains that implement HOTs about the qualities in somatosensory cortexes. We would entertain all of this, yet find we could still conceive that the creatures thereby described lacked awareness of the qualities in their brains. Conceivability entailing possibility, this means the failure of HOT panqualityism. Perhaps the theory can get qualities into the sort of order to match what we know, but it cannot account for our *awareness* of qualities. Note well: we should distinguish this argument from the standard zombie argument concerning *phenomenal consciousness*. That argument invokes zombies who lack sensory qualities *and* who lack also awareness of those qualities. In our case panqualityism guarantees that our material duplicates instantiate sensory qualities; what's allegedly missing is their awareness of these qualities. They are *awareness zombies*.

I will analyze and reject Chalmers's grounds for the claim that HOT panqualityism is threatened by awareness zombies. This leaves the positive motivations for the theory unobstructed.

ii. Why does Chalmers consider that panqualityist awareness zombies are conceivable? With standard physical zombies, who lack phenomenal consciousness, Chalmers emphasizes that the physical consists of *structure and dynamics*. One is thus prompted to conceive of a pure structure-and-dynamics

world, and sure enough finds that consciousness needn't be instantiated. In our case the bit of supplementary theory, aimed to get us conceiving in the right direction,⁷⁰ is "[a]*wareness involves phenomenology*, and there are good reasons to think that no mere functional state can constitute phenomenology . . . one can conceive of any such functional state in the absence of phenomenology, and in particular in the absence of awareness."⁷¹

The anti-HOT panqualityism zombie argument is thus to operate much like the anti-physicalist zombie argument—it hangs on a failure to reductively capture a certain phenomenology. In our case the missing target is narrower than *phenomenology in general*—the target of standard zombies. Chalmers apparently holds that, in addition to being aware of sensory qualities, we're *aware of our awareness* of sensory qualities. This further object of awareness—awareness itself—comes with its own patch of phenomenology: a qualitative feel. It's this feel which our HOT-panqualitative duplicates are alleged conceivably to lack.⁷²

It might seem an odd move to press the lack of a certain qualitative content against panqualityism. At this point in the dialectic, couldn't the panqualityist reply just by building the allegedly missing qualities into our, hence our HOTpanqualityist duplicates,' constitution? Not quite—for recall that the HOTpanqualityist position is that awareness is supplied by higher-order thought. It follows that if specific sensory qualities attach to awareness, these must be provided by the HOT component of panqualityism. And a HOT is avowedly a 'mere functional state.' So Chalmers's objection is well founded.⁷³

iii. I believe the objection fails, however. It's true that *if* awareness had phenomenology, then this, like sensory quality in general, would be hard to functionalize.⁷⁴ But I deny that mere awareness has phenomenology.⁷⁵

That awareness might lack phenomenology doesn't appear terribly surprising, when considered as a general matter: It is via awareness that we encounter sensory qualities and the appearances of things, but why should the faculty that presents sensory qualities to us itself make some appearance among our sensory qualities?⁷⁶ That would be akin to the camera appearing in the periphery of every shot of a television show.⁷⁷ It seems that there at least *could* be creatures for whom awareness contributed no sensory contents. For them, consciousness would be completely 'transparent' to its first-order objects. *Prima facie*, therefore, it is an open question whether we are such creatures. We must examine the evidence.

iv. In claiming that awareness lacks phenomenology, I deny that we are aware of awareness.⁷⁸ Were we aware of awareness, we could expect such second-order awareness to have phenomenology,⁷⁹ and Chalmers's strategy would be

vindicated. A lot hangs on this issue: Chalmers seems to concede that a deflationary account of awareness could combine with an 'informational' structuring of awareness to finesse such things as the structural mismatch problem.⁸⁰ His objection to this strategy rests on the alleged difficulty of analyzing awareness, which rests in turn on his claim that awareness has phenomenology.

Faced with the claim that awareness has phenomenology, a sensible approach is to search for its quality in consciousness. Kriegel posits a distinctive, pervasive and diffuse 'feel' contributed by awareness to our overall field of conscious sensory qualities.⁸¹ But he is no more specific than this, and I don't recognize the phenomenology in question from this description.⁸²

Chalmers refers us to some examples of the phenomenology of awareness which he employs elsewhere.⁸³ I understand, however, that he doesn't wish to hang too much on these particular cases, relying instead on what he sees as the plausibility of the general claim that we're aware of awareness. Accordingly I won't directly examine his examples; I'll only say that in each case where it is plausible that a distinctive phenomenology obtains, it's at least as plausible that it attaches (or is felt to attach) to the object of experience, or to things like qualia of mental effort, not to the manner of experiencing.

Some phenomenological reflection, for what it's worth (the difficulty of adjudicating this sort of dispute is inversely proportional with its proximity to the raw experience; indeed one might have imagined disagreement impossible at this distance, but quite the contrary). In being aware of *red*, I just don't know what my alleged awareness of my awareness of red is meant to feel like; I find only the redness. When you ask me to attend to the relational property of *my being aware of the redness*, still all I find is the redness—I don't seem to enter the picture (in respect of that redness). Of course I *know* I'm aware of redness, since *there it is* for me, subjectively. Similarly, I know there's a camera shooting a television scene, although I can't see the camera, only its output.⁸⁴

I think, strange as it may sound, that we *infer* that we're aware, because there are qualities present to us subjectively. One feels the qualities, but not that which goes into one's feeling them. This addresses the challenge sometimes leveled against those who reject the awareness of awareness: *If we're unaware of awareness, how could we possibly know that we're aware?*⁸⁵ This is a strange question, however. We who reject second-order awareness accept awareness. Awareness is the subjective presence to one of qualities. Now, since we accept awareness, we accept that qualities are subjectively present to individuals. Those individuals, in noting the qualities of which they're aware.⁸⁶

v. Those who allege a feel to awareness are not making things up: they surely detect *something*, phenomenologically. The question is what it is that they

detect. It seems a distinct possibility, for the skeptical, that they're misclassifying some more or less subtle feature (or features) of 'first-order' phenomenology. There are indeed some phenomenological factors with a tendency to confound. For instance, since one can infer that one is aware, there is the feel of the (pretty routine) conscious thought 'I am aware.' There is also the feel of 'self-awareness': the conscious sensory qualities associated with one's own body and mind (including the feels of prevailing emotional tenor, of bodily pains, of intentions, wishes, memories). Given these two items, it follows that one can *be aware of oneself as a thing that is aware*. Is there anything to the alleged sensory quality of awareness beyond this feeling? But this is not a phenomenology *of awareness*. It doesn't require awareness itself to be conscious, any more than *being aware of myself as a thing that is watching a show shot by TV cameras* requires me to see the cameras.⁸⁷

vi. There's also a serious question concerning what the sensory content of the alleged feel of awareness *could even be*. Proponents seem clear that this feel is an additional sensory content beyond the other qualities one is aware of (those pertaining to the environment, one's body, thoughts, etc.).⁸⁸ Now, we may ask, does this extra ingredient have its own, 'isolated,' feel—is it a standalone qualitative ingredient in consciousness; or, is its feel somehow interpenetrated by the other, first-order, qualities of which one is aware? Problems arise either way. If awareness has its own distinctive feel, which qualitatively speaking has nothing to do with, and makes no reference to, the other, first-order, qualities, then it is very hard to see how, in experiencing this quality, one could apprehend it as a feeling of *awareness of these (first-order) qualities*, i.e. as the very item it is supposed to be. Advocates of the phenomenology of awareness purport to identify it by its conscious feel. But this colorlessness would presumably have rendered the feel of *awareness of first-order qualities* unidentifiable as such, and likely wholly mysterious: a detached phenomenal UFO.

So it seems that the feel of awareness must somehow be *suffused* with the qualities that the awareness is of—the first-order qualities pertaining to experience of the environment, body, mind, and so on.⁸⁹ Then, at least, it exhibits to consciousness the intimate connection which it bears to these qualities. But now the position appears to be this: I am aware of a set of first-order qualities and aware, additionally, of an awareness-quale that *phenomenally includes* reference to (is 'stained by') these same first-order qualities. It seems to follow that I get every first-order quality *twice* in consciousness: once in its own right (as a 'floor-level item,' in Kriegel's phrase), and once more as 'staining' the feel of my awareness of all these first-order qualities. This duplication is unavoidable, since the sensory quality of awareness is posited as an item *additional* to the first-order qualities, while containing, in its feel (where else?), reference to

them. Yet, while I cannot locate the feel of awareness, I am certain that I don't have this doubling of qualities in my experience. The feel of awareness is here construed as a kind of mirror, giving reflection of every quality presented to it. I do not have this mirror, only its putative objects.⁹⁰ Either my phenomenology (or 'introspection') is atypical, or there's some confusion it seems in the doctrine of the feel of awareness.

vii. If awareness lacks phenomenology, as I have suggested, then there is not a distinctive kind of sensory quality that the HOT apparatus must contribute to consciousness. Awareness is as it were *behind the lens*.⁹¹ This means that, as against Chalmers, there is no phenomenological residue left unanalyzed by HOT panqualityism.⁹² While there is good reason to doubt that any sensory quality could be a purely functional affair, there is no such reason to doubt that awareness itself, unassociated with any sensory quality, could be a functional property. For all we presently know, higher-order thought may be the correct analysis.^{93,94}

10.5

i. I conclude that (1) the subject combination problem reveals a deep lack of theoretical motivation for panpsychism; (2) panqualityism, with the addition of a HOT apparatus for awareness, has the resources to make serious inroads into the nonsubject combination problems; and (3) Chalmers gives us no reason to doubt the adequacy of the sort of functional analysis of awareness offered by HOT theory.

ii. It seems that if constitutive Russellian positions enjoy significant advantages over mainstream physicalism and dualism,⁹⁵ then, given its advantages over panpsychism, panqualityism has a fair claim to be our best hope for a theory of consciousness.

Notes

- 1. See Chalmers this volume. Some of these problems have been knocking around for a long time. But it is fair to say that no one so far has separated and clarified them in such a comprehensive and careful way as Chalmers does.
- 2. My concern in this chapter is to contrast constitutive Russellian versions of panpsychism and panprotopsychism (see Chalmers (this volume) for these positions). I am with Chalmers in considering them the most promising variants of the general positions they represent (for reasons see again Chalmers's paper). I will henceforth generally refer to these variants simply as 'panpsychism' and 'panprotopsychism.'

- 3. I won't try to deflect here the sense some have that the very notion of unexperienced qualities is incoherent. See Coleman (2015; 2013; 2012) for efforts to make sense of unexperienced qualia.
- 4. See Chalmers (forthcoming) for an account of these advantages.
- 5. In another possible world, I should have liked to call panqualityism plain 'physicalism.' But since our actual physicalists for some reason feel the need to deny that basic material entities have any qualities at all-i.e., not-merely-relational properties about which it is appropriate to say there is something it is like, e.g., colors—I am forced into a fruity name (Chalmers gets it from Feigl (1958), who credits in turn S. C. Pepper). These same actual physicalists spend much of their time adopting theoretical contortions of various painful sorts, in order to evade the glaring truth that if the ingredients that compose us (and the world) have no qualities, there is no possibility of our experiencing qualities. The confident counterfactual physicalists with whom I would wish to associate are by now busily working out a fullyfledged qualitative physics, which explains not only the material dynamics and development of (what we understand as) physical systems but also the dynamics and development of (what we understand as) mental systems, all from a single set of basic principles. They can explain why pains have the causal profiles they do (and must have), and possess a deductive explanation of how paracetamol cures headaches. Meanwhile, their actual counterparts are exploring hopefully the hypothesis that there are no qualities only wholly deceptive representations of qualities within our deluded minds. The contrast, when once stated, is damning.
- 6. I'm using 'micromentality' to cover both microexperiences, as on panpsychism, and microqualities (which are nonphenomenal), as on panprotopsychism. In fact panprotopsychists often deny that microqualities are mental, on the ground that they are not conscious. I'm suppressing that point for ease of exposition.
- 7. See e.g. Rosenthal 2005.
- 8. Block (2011) claims it offers inconsistent conditions on a conscious state.
- 9. Another 'bridging' objection to HOT theory runs that infants lack the architecture for HOTs, though they are plausibly conscious.
- 10. This is the *Russellianism* in constitutive Russellian panpsychism/panprotopsychism.
- 11. On the Nature of Things, book II: 849–55.
- 12. To be clear, if, as in *fusion*, the ingredients—things, properties—do not survive production of the novel entity, then this is causal emergence, not constitution. See Chalmers's discussion (this volume) of Seager's 'combinatorial infusion' as an option for panpsychists.
- 13. Lockwood and Foster have recently been concerned with this problem. Lockwood: "What, one may ask, is the use of attributing, say, embryonic color to the ultimate physical components involved in the neuronal goings-on that are supposed to be constitutive of a phenomenal patch of red, if these self-same constituents are also to be capable of figuring in auditory or olfactory experiences which are wholly devoid of visual phenomenology?" (Lockwood 1993, 277). Foster (1991,127): "How ... could a different arrangement of pain-particles yield a visual experience or a surge of anger?" And, "if we are dealing with a visual experience, then presumably we have to assign visual qualities to the constituents of the neural item in order to account for its introspectible character. But these qualities would not be appropriate to the roles of similar physical constituents in neural items correlated with non-visual experiences" (Foster 1991, 129).
- 14. These would still have to show up in their products, and some would surely seem qualitatively too far away for this to be possible.
- 15. It's pretty clearly unacceptable to assign fixed qualities to ultimates in isolation, and then say that in combination they simply *lose* these (this would be to imagine that a red and a white ultimate each independently 'turns' pink upon meeting). For, as Lucretius says: "if they were to give up from their bodies, Their own power of feeling, and acquire another one, What was the point of giving them in the first place, What is taken away?" (*On the Nature of Things*, book II: 924–27). Generally speaking, fundamental intrinsic properties must (1) *remain in play in constitution* and (2) *be directly explanatorily relevant to their products*. This point later forms the core of my objection to panpsychism.
- 16. Recognition of something like these problems is probably behind Feigl's suggestion that on panqualityism the fundamental qualities had better be relatively "colorless" (Feigl

1971, 308). But then: (1) We'd face a problem around their having enough qualitative "oomph" to generate macroqualities at all. (2) Making the basic qualities homogenous ("colorless") doesn't apparently help with the problem of how we get, *via* their combination, to such remote and distinctive locations in quality space as we actually find.

- 17. Chalmers says "given a Russellian view, it is not at all easy to see how these [micro]structures could be [allowed to be] so different that they yield the vast differences between macrophysical and macrophenomenal structure" (Chalmers this volume; my emphasis).
- Sellars's original grain problem is more plausibly about bridging: *neurons* are discontinuous, he says, while the experiential field is continuous, yet the latter is supposedly constituted by the former (See e.g. Sellars 1963, 35).
- 19. Our focus is constitutive Russellian versions of panpsychism and panprotopsychism, recall.
- 20. The very famous passage being the one occurring at page 160 of his 1890. See Strawson (2006), and accompanying papers, for more recent discussion of the subject problem.
- 21. On the Nature of Things, book II: 920-23.
- 22. For the definitive version of this argument see Goff (2009).
- 23. See Coleman 2013.
- 24. Taken from Chalmers (this volume), who credits Goff 2009. Note that in Chalmers's formulation 'panpsychism' reads 'constitutive panpsychism'; this is unnecessary for us given the earlier decision to focus on constitutive Russellian positions.
- 25. This isn't to say there aren't other explanatory roles played by fundamental posits—but such properties must *at least* explain their higher-level instances, where they have such.
- 26. See Goff this volume.
- 27. This sort of thing has been suggested to me by Pat Lewtas.
- 28. Chalmers this volume.
- 29. Chalmers this volume.
- 30. Jill is *taller than* Bob thanks to their intrinsic properties; and so on.
- 31. In respect of their subjectivity, at least. People arguably form into larger entities such as crowds, senates, and nations. But, as James notes (just after his famous passage about combining feelings) these alleged entities do not have a 'group mind' in any serious sense.
- 32. With panpsychism and its microsubjects now out of the way, it's worth recording, as regards the broader argumentative context, that regular physicalism faces all these combination problems (the question of how qualities combine must be faced sooner or later, at micro- or macrolevels) *plus* the problem (surely insurmountable—see Coleman 2015) of manufacturing qualities out of the nonqualitative. It follows that even if we can make no decent progress here with panqualityism's problems, it is in at least as good shape as regular physicalism; really much better shape, once one takes in the irreducibility of qualities.
- 33. See Chalmers this volume.
- 34. This provides a neat means of removing apparent action at a distance (not to mention apparently instantaneous—so faster than light—effects!) when entangled particles are very far away one from one another.
- 35. Chalmers seems to see separate instantiation as a bar to combination. We may remove this via entanglement, without going so far as to say that entanglement is *all it takes* for such combination. We may want to keep entanglement (or co-location) as a necessary but not sufficient condition, so as to avoid anything that is entangled instantiating only a single quality—for instance the entire universe, if this is wholly entangled with itself.
- 36. Chalmers sees the possible utility of quantum holism for panpsychists like this: "If subjects' experiential fields could be *identified with* physical quantum wholes—large physical fundamental states—this might remove the need to account for them combinatorially" (Chalmers this volume). Though not a combinatorial panpsychism, Chalmers still counts this a constitutive Russellian variety (see Chalmers this volume).
- 37. All portions of matter would then have all been entangled from the start, plausibly remaining so no matter how far apart everything subsequently drifted.
- 38. This challenge aside, I think William James revealed cosmo-panpsychism as incoherent. Assuming our reality as subjects, we are on this view phenomenal components of the

universe-subject. It follows that the universe-subject is conscious of all the things you and I are conscious of. Yet I can, for example, *sincerely wonder* what you're thinking. The universe-subject cannot *sincerely* wonder what you're thinking, though, since it already knows (by feeling) what you're thinking. So it seems my sincere feeling of wonderment cannot, after all, be a phenomenal component of the universe-subject, as against the initial supposition. I believe this clever argument is in *A Pluralistic Universe* somewhere, though I can no longer locate it.

- 39. Chalmers this volume.
- 40. Lee Smolin has suggested that the varying qualities we experience correspond to energy fluctuations: in a panqualityist context, this suggestion can naturally be expanded to take in the whole universe, with the fluctuations construed as its aspects.
- 41. This apparently clashes with Chalmers's reading of entanglement, or at least, the one he offers the quantum holist panpsychist, whereby the new state wholly supersedes the unentangled elements. I'm more inclined to construe the latter as helping to constitute the new entity. A physicist friend tells me that (for example) entangled electrons survive in the new state somewhat 'like sausages and potatoes in toad-in-the-hole.' This makes sense, though I didn't know that toad-in-the-hole had potatoes. Thanks to Paul Cook for discussion.
- 42. Which neatly parallels—even helps us perhaps visualize—corresponding microphysical operations.
- 43. Or something like Bohm's idea (1980) of a much finer-grained fundament of entities and properties beneath the quantum mechanical level currently considered basic.
- 44. Again Smolin's suggestion about quality correlating with energy fluctuations comes to mind.
- 45. On string theory Lockwood says, "it seems incomprehensible that different combinations of collective or individual string states could generate the qualitative diversity that is manifest at the phenomenal level. It seems inconceivable in much the same way, and for much the same reasons, that it is inconceivable that an artist, however skilled, should conjure the simulacrum of a Turner sunset from a palette containing only black and white paints" (Lockwood 1993, 276). This remark is puzzling from the standpoint of a mere worry about the *number* of slots available in the microphysical realm for qualities, given the range in which strings can vibrate. This suggests that the real worry as regards the micropalette and 'diversity' concerns *incommensurability*. See below.
- 46. See Hartshorne 1934, especially the first half. Of particular note are (i) his argument that it's possible for an auditory quale to be qualitatively closer to a visual quale than to another auditory quale (he compares a fife note to *silver* and to a *dull thump*) and (ii) his clever discussion of synaesthesia as evidencing the continuum hypothesis.
- 47. Objection: If colors are like electric qualia, but color qualia are unlike taste qualia, then electric qualia cannot be like taste qualia. Yet we know that transitivity of similarity fails even among the colors.
- 48. Though the ultimate suggestion is of course that we drop the modality-based conceptual scheme if possible.
- 49. Although it may be that some of the macroqualities we know of are *also* fundamental—but which ones and how to tell? Work to be done. I've written a little more about the deductive project mentioned here in Coleman 2015, but we await the genius.
- 50. Thus I don't find the objections to HOT theory mentioned in section I persuasive: no room to explain why, however. Rosenthal is also undaunted by the objections.
- 51. There are theories which make the higher-order representation more like perception than thought (see e.g., Lycan 2004). There's also a self-representational view like Kriegel's (2009), where the conscious state and the state that provides awareness of it are more tightly bound—into the same mereological complex in fact. I prefer HOT theory for reasons explained elsewhere.
- 52. See the next three subsections for a wrinkle on this claim.
- 53. See Bayne and Chalmers 2003.
- 54. James (1890, 158-59): "the sum [of experienced qualities] itself exists only for a bystander who happens to overlook the units and to apprehend the sum as such." The bystander I propose is a HOT. Being appropriately related to a suitable HOT could perhaps be

understood as a panqualityist version of a phenomenal bonding relation (after Goff 2009 and Chalmers this volume).

- 55. See Kriegel (2009, ch. 7) for well-informed speculation about the kind of neurological integration likely required.
- 56. Though Sellars (1963, 37) seems open to the idea that ultimate homogeneity might by itself be enough to finesse the grain problem. In this case, the appeal to HOT theory is not needed. It will still find gainful employment helping with structural mismatch, however (next subsection).
- 57. Rosenthal 1999, 345.
- 58. This move becomes tempting to more mainstream materialists, once the idea arises of a mediating representational mechanism in introspection or consciousness. The first half of Pereboom's 2011, for instance, toys with the idea that our phenomenal concepts are *wholly* deceptive, representing to us in introspection qualities that nowhere obtain. Yet how the content of such representations could be supplied is left unexplained. Against this view I juxtapose Sellars (1963, 30): "we have taken them [qualities] out of our world picture altogether. We will have made it unintelligible how things could even *appear* to be coloured." See also Coleman 2015.
- 59. Or it may be that, taking note of the earlier model of qualitative combination, the constituting qualities of a given node exist now only as contributions to a whole which has taken the metaphysical upper hand. In this case, while the whole is in being, the composing qualities are only implicit within it, as regards their original form, and are not there literally to be observed or experienced, in this form. I leave this thought hanging, as it is not clear to me. It has obvious connection to Sellars's line of thought in note 56.
- 60. Lockwood 1993, 278. See Foster's (1991, 127) talk of 'distance' from what is experienced. Foster doesn't claim that taking such a distance is impossible, however; his main difficulty for Russellianism is what I have called the *incommensurability problem*.
- 61. This is Lockwood's analogy, which he claims does not carry over to the mental case (see Lockwood 1993, 277–78).
- 62. Lockwood 1993, 274. Chalmers says "the macrophenomenal structure of my visual field is prima facie very different from the macrophysical structure of my brain" (Chalmers this volume).
- 63. Feigl also discusses this sort of proposal, attributing it to Carnap: "it would have to be assumed that one area of the cortex 'taps' or 'scans' other areas....Likewise, one would have to assume that the effect in the second [scanning] area reflects only certain gross features of the intricate and multifarious process patterns in the first... the second area...corresponds to the sensing of raw feels" (Feigl 1958/1967, 91).
- 64. Lockwood 1993, 275.
- 65. Lockwood arguably leaves consciousness open to such a flanking manoeuver: he says functionalism must fail "at least if put forward as a *global* theory of mind" (Lockwood 1993, 275; my emphasis).
- 66. Chalmers (this volume) also thinks the structure of the modalities is a source of structural mismatch, but I've rejected the ultimate reality of that structure: I think it might well dissolve upon ideal refection.
- 67. Chalmers identifies another troubling aspect of the structural mismatch problem: the notion that, given Russellianism, qualitative and physical *property structures* would have to match up. For example, "if mass has a scalar structure, the associated [quality] (what plays the mass role) has a scalar structure. If charge has a binary structure, the associated [quality] (what plays the charge role) has a binary structure" (Chalmers, this volume). I have to admit to being bamboozled by this impressive difficulty. If there is a manifest structural clash here—at the moment I struggle to see whether there is or not—then my obvious remedy is to invoke the HOT apparatus again. Perhaps this ensures that the property structures we *experience as* belonging to qualities are somewhat artificial (the product of the HOT filtering process in preparing a percept). Unscreened, qualities perhaps have the structural properties of physical properties. Whether anything like this would work I don't know. This is a very interesting problem indeed. I wonder how the matter would appear if we tried to match microqualities to

individual vibration states of strings: would that produce any structural clash? Like a string, I wave in the breeze here.

- 68. It may be tempting, but it would be no less question-begging, to say the relation is *constitu-tion*: that the ultimates you are conscious of are the ones that constitute your conscious mind. The whole question is in virtue of what do these conscious ultimates not also constitute mine. Goffian phenomenal bonding, at least as presently understood, is of no help either: to say a certain set of ultimates are phenomenally bonded is to say that they are the ones co-conscious for some subject; this is just to *describe* the state of affairs that requires *explaining* in independent terms.
- 69. Lockwood maintains that "there are no distinctively introspective meta-mental representations . . . whose separation from their mental objects could help us resolve . . . the grain problem" (Lockwood 1993, 278). Aside from the reference to introspection— HOTs do their representation in 'first-order' consciousness—this is the gist of the present HOT-based proposal: I've no idea why Lockwood rules it out. His claim appears to be an empirical one.
- 70. I'm not saying there's anything wrong with such prompting: to set up a zombie argument properly, *explanandum* and *explanans* must be made precise enough for the thinker to conceive informatively. If we are not told to conceive of the physical in terms of structure and dynamics, our conception of the physical is left too open-ended. That's plausibly why some people react to the standard zombie argument by saying it's inconceivable that all the physical stuff could be there without consciousness, since consciousness just *is* part of the physical stuff: they haven't had (or heeded) the prompt about structure and dynamics. It's very hard to imagine that consciousness *just is* structure and dynamics. Note that here the prompting primarily concerns the nature of the *explanandum* (awareness, that it has phenomenology) whereas with the standard zombie argument it primarily concerns the *explanans* (the nature of the physical).
- 71. Chalmers this volume, my emphasis.
- 72. Interestingly, then, this argument doesn't seem directly to concern the irreducibility of awareness *as such*. Rather, the idea is that awareness necessarily comes with a phenomenal complement. Since the complement is unanalyzable, awareness must also be unanalyzable, as the complement is bound up with it. See also the next note.
- 73. Still, one might wonder just why Chalmers proceeds via phenomenology in this way. Why does he not simply claim directly that awareness isn't functionalizable? The answer is that zombie scenarios depend on there being a sensory quality 'toggle' between the actual world and putative zombie world. One conceives of the relevant 'zombified' property by conceiving of the absence of its associated sensory qualities (e.g., zombie water is H₂O without waterish sensory qualities, zombie heat is MKE without heatish sensory qualities; the thrust of the standard zombie argument is that removal of its associated sensory qualities amounts to removal of the very property in the case of phenomenal consciousness). It's not possible to construct a zombie scenario without framing the target or *explanandum* in sensory quality terms, in fact (this is also true of the explanans, but showing why that is would take us too far afield). It follows that any item not associated with a set of sensory qualities is not a valid target for a zombie argument. Setting aside its structural impact on the experiential field, which should be deducible from the brain's functional structure, awareness is one such item, I argue momentarily. However, this excursion into the general mechanics of zombie arguments is not needed to block Chalmers: his present anti-panqualityist argument depends on the explicit claim that awareness has phenomenology; in the interesting (i.e. not-merely-structuring) sense, I deny this.
- 74. In fact my diagnosis of the standard zombie argument is that it depends much more on the elusiveness of qualities to functionalization, than it does on the elusiveness of sheer consciousness (aka awareness). There is evidence for this in how the argument is sometimes put. Churchland (2014, 37), for example, describes zombies as our physical duplicates, "whose subjective *qualitative* mental life is simply absent"—significantly, he doesn't mention consciousness at all in setting up the zombie challenge to physicalism. This point is implicit in Byrne (2006). See also previous note.

75. In the sense of being associated with sensory qualities—it has no 'feel.' We did admit (3: viii-ix) an *impact* of the HOT mechanism on our sensory qualities—for it has a 'smoothing' and structuring effect. But these structural aspects of our phenomenology, I maintain, ought to be recoverable from details of the HOT system's interaction with its sensorily qualitative target among our brain tissues—these explanations would be on a par with those which might be given of the boundedness and unity of our HOT panqualityist duplicates' experiential fields, also via the HOT mechanism. While the functional HOT mechanism cannot *manufacture* sensory qualities, as Chalmers rightly points out, it is able effectively to *corral* and to *filter* existing sensory quality instances (to use a metaphor, though HOTs can mold the clay they're given, they cannot produce the qualitative clay itself). We thus do not get any zombie-susceptible aspects of phenomenology.

N.b.: If he did not (at least implicitly) accept that the HOT mechanism could be a priori connected to the structure of phenomenology, it seems that Chalmers could not have run the objection we encountered when discussing phenomenal unity: that multiple HOTs would prevent the unity of consciousness. In that objection, he infers from HOT-structure to phenomenological structure.

- 76. This is not the denial that awareness is real (the view Chalmers ascribes to James in his radical empiricist phase): of course we are aware of qualities. But we are, in the relevant sense (see fn. 86), unaware of our awareness of qualities.
- 77. Or the eyes appearing in the visual field—cf. Wittgenstein in *Tractatus*.
- 78. Lockwood says, "sensory phenomenology belongs, so to speak, to that tip of the neurophysiological iceberg *which projects above the surface of awareness*. We are to regard it as a part or aspect of the reality of the brain that is directly present to the conscious mind" (Lockwood 1993, 282; my emphasis). My position is implicit here: sensory qualities project above the surface of awareness, but awareness itself need not so project. Strictly, we are not even aware of awareness (i.e., the HOT mechanism) as structuring our sensory field: for being aware only of the *effects* of some item is not the same as being aware of that item. As Kriegel says, something that structures phenomenology "makes a difference to the phenomenology—without being an *item in* it" (Kriegel 2009, 172).
- 79. It seems that if we are aware of some *x*, then *x* is like something for us, in the Nagelian sense. That means in turn that *x* is associated with certain sensory qualities. How could we be aware of *x*—in the sense relevant to consciousness—without *x* being like something for us?
- 80. See Chalmers this volume.
- 81. See Kriegel 2009. Similarly, Chalmers talks of a phenomenal "background acquaintance with our awareness" (Chalmers 2013, 5).
- 82. Gennaro (e.g., 2008) cannot locate the phenomenology of awareness either. See Kriegel (2009, ch. 5) for an ingenious explanation of why the phenomenology is elusive to introspection. This explanation, however, does not help with the fact (as I see it) that the phenomenology is not felt in a first-order way, as it had better be. We don't—can't—require introspection to confirm the presence of all the qualities we're aware of. Is it only by introspecting my experience that I know I sometimes see blue? This seems an unnecessarily technical requirement to verify my occasional awareness of blue. And this, that introspection is required to confirm the existence of a sensory quality, is not something Kriegel can anyway say: for since the feel of awareness is said to elude introspection, yet Kriegel claims to feel this feel, these must be phenomenological data outside of introspection he's relying on. I don't have those data. All of which indicates, additionally, in my view, that something is seriously wrong with the prevalent talk of 'introspection'—I have come to lose more or less entirely my grip on what this operation is supposed, phenomenologically and mechanically, to comprise—is it much more than staring very hard at a wall while mentally muttering the inanity 'This is an experience'?
- 83. In his reply to Hellie's commentary on Chalmers's *The Character of Consciousness* (Chalmers 2013).
- 84. Or I know there's an eye because the visual field is apparent to me. In fact the relationship is tighter in the case of awareness than in these examples: if I am dreaming then the visual
experiences I have do not (at that time) require an eye, at least not in the normal way. But if sensory qualities are apparent to me *at all*, that is because of my awareness.

- 85. See e.g. Kriegel (2009, ch. 4.3) for this sort of challenge.
- 86. In Dretske's terms (see e.g., Dretske 1999), we're *aware that* we're aware, but not *aware* of awareness—this is the distinction between fact awareness and object/property awareness. But mere fact awareness isn't what Chalmers has in mind in alleging a phenomenology of awareness—Dretske is clear that this kind of awareness is a phenomenology-free affair, in the sense that we can become that-aware concerning some fact just by reading about it. The *fact* about which we're thereby aware need communicate nothing phenomenologically. The (property) awareness of sensory qualities which we are (fact) aware (i.e., know) *that* we possess need therefore involve no qualitative feel itself.
- 87. Or, perhaps more simply, any more than *being aware of myself as a thing that is watching* requires me to see my eyes.

Objection: Yet we do sometimes see eyes and TV cameras, and that's how we know they're there; do we sometimes, then, become aware of awareness (in 'introspection')?

Reply: Because awareness is so tightly connected with qualities being subjectively there for one *at all*, there's no need, in this case, to observe the organ 'from the outside.' I really doubt there's much to introspection, beyond, perhaps, thinking about the qualities one is aware of. But any sensory (or cognitive) qualities accruing to *such* thoughts, beyond the qualities of their sensory objects, are not contents the HOT apparatus is charged with generating, so no objection to HOT panqualityism lies in this direction. The HOT component's job is to produce first-order awareness. It follows that I deny, against Rosenthal, that HOTs can take other HOTs as objects (this is Rosenthal's model of introspection).

In fact there is not even an objection here that might affect the reply to Chalmers. If we *could* be introspectively aware of awareness (which I deny), that would involve a HOT targeting another HOT, which was targeting in turn certain 'floor-level' sensory qualities. If 'awareness' contributed sensory qualities to consciousness here, those would be qualities pertaining to the 'lower' HOT, the target of the meta-HOT. But of course HOTs *do* have qualities, because they are simply neural items, and on panqualityism all neural items (all items in fact) are constituted of qualities. These qualities are not (*per impossibile*) *produced by* a HOT's functional aspect, they rather *realize* or carry that very aspect (compare: a mousetrap is not of this configuration of wood *because* it's a mousetrap, but vice-versa). So there would be no difficulty with a HOT contributing qualities to consciousness, were it the object of a further HOT whose functional property made the former HOT conscious. I simply deny, on phenomenological grounds, that this occurs, however.

- 88. See e.g. Kriegel 2009, 180. It seems the feel of awareness couldn't very well be an 'aspect' independently added to each first-order quality of which one is aware: then in seeing a red rose one would have 'red-rose-plus-my-awareness' phenomenology; but this is not how red roses appear: they just appear red and rose-y. The popular doctrine of the transparency of experience (see e.g. Harman 1990) could never have got up and running, were the feel of awareness an aspect of every first-order sensory content.
- 89. This is in fact the line Kriegel (2009, 180) takes.
- 90. Even if it's in the phenomenological background, as Kriegel says, a faint mirroring or duplication there must nonetheless (phenomenally) be.
- 91. This metaphorical location permits it of course its structuring effect on phenomenology. Another plausible example of such a phenomenon is *memory*. Memory is not behind the lens but *below the surface* as regards appearances. Like awareness, all we get from memory is (delivery of) some items of which we are aware, but memory itself makes no appearance in terms of sensory content. To be sure, there is a feel to *trying to remember*, also a feeling of *having successfully recalled*, but the memory process itself—what comes in between these conscious events—is wholly obscure to awareness.
- 92. For I claim that the structuring aspects of the HOT system will be deducible—*given* details of the qualitative clay on which it is to effect its molding, naturally.
- 93. Objection: If there is no phenomenology of awareness to be theoretically analyzed, how could the earlier critique of panpsychism rest on its inadequacy as a theoretical analysis of macrosubjectivity? Cannot panpsychism evade critique in the same way HOT panqualityism does?

Reply: Panpsychism's failure really consisted in an inability to say anything useful about the *constitution* and *structure* of macrosubjects, in particular, about why they have the unity and boundaries they do. By contrast, one of HOT panqualityism's strengths is the explanation it can provide of these features. The accusation against panpsychism was never that it could not account for awareness *as such*, since this it simply presupposed!

- 94. It will likely be said that it remains perfectly conceivable that our HOT-panqualityist duplicates might lack awareness. But as awareness has no proprietary sensory quality associated with it, I'm simply unsure what someone could be conceiving of who made this claim, since zombie-style conceiving requires a sensory quality toggle between the actual world and the relevant zombie world—some qualitative content that we can subtract, in conception, from the zombie world. In this sense, 'awareness zombies' are inconceivable.
- 95. As Chalmers (forthcoming) argues.

Bibliography

- Block, Ned. "The Higher-Order Approach to Consciousness Is Defunct." *Analysis* 71.3 (2011): 419–31.
- Bohm, David. *Wholeness and the Implicate Order*. London: Routledge and Kegan Paul, 1980. Rpt. by Routledge Classics, 2002.
- Byrne, Alex. "Color and the Mind-Body Problem." Dialectica 60.2 (2006): 223-44.
- Chalmers, David. "The Contents of Consciousness: Reply to Hellie, Peacocke, and Siegel." Analysis 73.2 (2013): 345–68.
- Chalmers, David. "The Combination Problem for Panpsychism." *Panpsychism*. Ed. Godehard Brüntrup and Ludwig Jaskolla. Oxford: Oxford University Press, this volume.
- Chalmers, David. "Panpsychism and Panprotopsychism." *Russellian Monism*. Ed. Torin Alter and Yujin Nagasawa. Oxford: Oxford University Press, *forthcoming*.
- Chalmers, David, and Tim Bayne. "What Is the Unity of Consciousness?" *The Unity of Consciousness: Binding, Integration, Dissociation.* Ed. Axel Cleeremans. Oxford: Oxford University Press, 2003. Ch. 1.
- Churchland, Paul. "Consciousness and the Introspection of 'Qualitative Simples." Ed. Richard Brown. Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience. Dordrecht: Springer, 2014. 35-56.
- Coleman, Sam. "Mental Chemistry: Combination for Panpsychists." Dialectica 66 (2012): 137–66.
- Coleman, Sam. "The Real Combination Problem: Panpsychism, Micro-Subjects, and Emergence." *Erkenntnis* 79.1 (2013): 19–44.
- Coleman, Sam. "Neurocosmology." *Phenomenal Qualities: Sense, Perception and Consciousness.* Ed. Paul Coates and Sam Coleman. Oxford: Oxford University Press, 2015. 66–102.
- Dretske, Fred. "The Mind's Awareness of Itself." Philosophical Studies 95.1/2 (1999): 103-24.
- Feigl, Herbert. "The 'Mental' and the 'Physical." Minnesota Studies in the Philosophy of Science 2 (1958): 370-497. Rpt. (with a postscript) as The "Mental" and the "Physical." Minnesota: University of Minnesota Press, 1967.
- Feigl, Herbert. "Some Crucial Issues of Mind-Body Monism." Synthese 3.4 (1971): 295–312.
- Foster, John. The Immaterial Self: A Defense of the Cartesian Dualist Conception of Mind. London: Routledge, 1991.
- Gennaro, Rocco. "Representationalism, Peripheral Awareness, and the Transparency of Experience." *Philosophical Studies* 139.1 (2008): 39–56.
- Goff, Philip. "Why Panpsychism Doesn't Help Us Explain Consciousness." *Dialectica* 63.3 (2009): 289–311.
- Harman, Gilbert. "The Intrinsic Quality of Experience." *Philosophical Perspectives* 4 (1990): 31–52.

Hartshorne, Charles. The Philosophy and Psychology of Sensation. Chicago: University of Chicago Press, 1934.

James, William. The Principles of Psychology. New York: Dover, 1890.

- Kriegel, Uriah. Subjective Consciousness: A Self-Representational Theory. New York: Oxford University Press, 2009.
- Lockwood, Michael. "The Grain Problem." *Objections to Physicalism*. Ed. Howard Robinson. Oxford: Oxford University Press, 1993. 271–92.
- Lucretius. On the Nature of Things. Oxford World's Classics. Translated by Ronald Melville. With an Introduction and Notes by Don and Peta Fowler. Oxford: Oxford University Press, 1997
- Lycan, William. "The Superiority of Hop to HOT." *Higher-Order Theories of Consciousness: An Anthology*. Ed. Rocco Gennaro. Amsterdam: John Benjamins, 2004.
- Pereboom, Derk. Consciousness and the Prospects of Physicalism. New York: Oxford University Press, 2011.
- Rosenthal, David. "Sensory Quality and the Relocation Story." *Philosophical Topics* 26.1/2 (1999): 321–50.
- Rosenthal, David. Consciousness and Mind. Oxford: Clarendon Press, 2005.
- Sellars, Wilfrid. "Philosophy and the Scientific Image of Man." *Science, Perception and Reality.* Ed. Wilfrid Sellars. London: Routledge and Kegan Paul, 1963.
- Strawson, Galen. "Panpsychism? Reply to Commentators with a Celebration of Descartes." Journal of Consciousness Studies 13.10–11 (2006): 184–280.

The Phenomenal Bonding Solution to the Combination Problem

PHILIP GOFF

11.1 The High Probability of Panpsychism

Panpsychism, the view that fundamental physical entities are conscious, is a highly probable theory of the natural world. Appreciation of this requires little more than getting our epistemic situation right.

Physics tells us much about the dispositions of fundamental natural entities but leaves us completely in the dark about their categorical nature. In knowing that an electron has a certain amount of mass, we know how it is disposed to resist acceleration and attract other things with mass. In knowing that an electron has negative charge, we know that it is disposed to repel other things with negative charge and attract things with positive charge. Everything natural science has to tell us about electrons concerns their behavior; we learn nothing about what an electron *is* independently of what it *does*. The only thing we know for certain about the categorical nature of natural entities is that at least some of them, for example you and I, are conscious.

We now have a theoretical choice. We can either suppose that the categorical nature of fundamental particles, such as electrons and quarks, is constituted of some form of consciousness, or we can suppose that they have some entirely unknown categorical nature. On the former supposition, the nature of macroscopic things is continuous with the nature of microscopic things. The latter supposition in contrast adds complexity, discontinuity, and mystery. The theoretical imperative to form as simple and unified a view as is consistent with the data leads us quite straightforwardly in the direction of panpsychism.

The main objection one comes across to panpsychism is that it is 'crazy' and 'just obviously wrong,' It is thought to be highly counterintuitive to suppose that there is something that it is like to be an electron, and this is taken to be a very strong reason to doubt the truth of panpsychism. But the view that time slows down at high speeds, that particles have determinate positions only when measured, that the Earth goes round the sun, or that our ancestors were apes were (indeed still are) also highly counterintuitive, to many 'just obviously wrong.' And yet the counter-commonsensicality of these views gives us little or no reason to think them false. It is hard to see why the fact that most Westerners living today happen to be pre-theoretically inclined to think panpsychism false constitutes a reason to think that it is false.

Probably the willingness of contemporary philosophers to accept special relativity, natural selection, and quantum mechanics, despite their strangeness from the point of view of pretheoretical common sense, is a reflection of their respect for the scientific method. We are prepared to modify our view of the world if we take there to be good scientific reason to do so. But in the absence of hard experimental proof, philosophers are reluctant to attribute consciousness to electrons.

However, whilst there is no observational data that supports panpsychism, there is a hard datum which counts in its favor: the existence of consciousness. The reality of consciousness is more evident to us than any empirical postulation. The existence of consciousness does not entail the truth of panpsychism, but it counts in its favor in the sense that panpsychism is the most unified picture of the world that is consistent both with its existence and with our observational knowledge. Compare the datum that the speed of light is measured to be the same in all frames of reference. This datum does not entail the truth of special relativity, but it counts in its favor in the sense that special relativity is the most elegant picture of the world consistent with it. The evident existence of consciousness supports the truth of panpsychism in much the same way that measurements of light support special relativity.

Whilst in the mind-set of thinking that physics is on its way to giving a complete picture of the fundamental nature of reality, panpsychism seems improbable, as physics does not attribute conscious states to fundamental particles. But once we realize that physics leaves us completely in the dark about the categorical nature of the entities it talks about, and indeed that the only thing we know for certain about the categorical nature of the universe is that some of it is taken up with consciousness, things look very different. All we get from physics is this big black-and-white abstract structure, which we metaphysicians must somehow color in with real categorical nature. Assuming the falsity of substance dualism, we know how to color in one bit of it: the brains of organisms are colored in with consciousness. How to color in the rest? The most elegant, simple, sensible option is to color in the rest of the world with the same pen.¹

11.2 Introducing the Combination Problem

Despite its obvious attractions, panpsychism suffers from a serious problem: the so-called combination problem, It is natural to suppose that my mind, the subject of my consciousness, is not a microscopic entity. Assuming the falsity of substance dualism, my mind is a macroscopic entity that derives its nature from the microscopic entities which compose it, ultimately from the entities that fundamental physics talks about, which the panpsychist takes to be conscious subjects. Somehow little subjects, such as electrons and quarks, come together to produce big conscious subjects, such as human brains. The combination problem is given by the fact that it's hard to make sense of this kind of combination.²

The inspiration for the combination problem is the following much quoted passage from William James:

Take a hundred of them [feelings], shuffle them and pack them as close together as you can (whatever that may mean); still each remains the same feeling it always was, shut in its own skin, windowless, ignorant of what the other feelings are and mean. There would be a hundred-and-first-feeling there, if, when a group or series of such feelings where set up, a consciousness *belonging to the group as such* should emerge. And this 101st feeling would be a totally new fact; the 100 feelings might, by a curious physical law, be a signal for its *creation*, when they came together; but they would have no substantial identity with it, not it with them, and one could never deduce the one from the others, nor (in any intelligible sense) say that they *evolved* it.³

Many philosophers, under the influence of this passage, claim to find some special conceptual difficulty in the idea of *feelings* or *subjects* combining. In fact, closer examination of the text surrounding this passage reveals that James's resistance to the summing of mental entities is grounded in a general resistance to the idea of *anything* combining:

[N] o possible number of entities (call them as you like, whether forces, material particles, or mental elements) can sum *themselves* together. Each remains, in the sum, what it always was; and the sum itself exists only *for a bystander* who happens to overlook the units and to apprehend the sum as such; or else it exists in the shape of some other *effect* on an entity external to the sum itself. Let it not be objected that H_2 and O combine of themselves into 'water,' and thenceforward exhibit new properties. They

do not. The 'water' is just the old atoms in the new position, H-O-H; the 'new properties' are just their combined *effects*, when in this position, upon external media, such as our sense-organs and the various reagents on which water may exert its properties and be known.⁴

Without much argument, James takes it to be evident that entities 'combine' only in the sense that their acting in concert gives rise to some distinctive perception in observers. Combinations exist only in the eye of the beholder. The 'combination problem' according to James goes as follows:

- 1. There are no combinations.
- 2. Therefore, there are no mental combinations.⁵

However, the more suggestive elements of the former paragraph have made more of an impression on contemporary philosophers than the argument of the latter paragraph. There does seem to be some deep difficulty making sense of distinct subjects combining to produce a greater subject; a difficulty we don't seem to face making sense of distinct objects in space combining.

In earlier work, I said the following:

Small objects with certain shapes, e.g. lego bricks, can constitute a larger object with a different shape, e.g. a lego tower. But it is difficult to see how, say, seven subjects of experience, each of which has a visual experience as of seeing one of the colours of the spectrum, could constitute a distinct subject of experience having a visual experience as of seeing white....Take the case of seven lego cubes placed on top of each other to make a rectangular tower. The mere existence of those bricks, each having a specific shape and location, necessitates the existence of the tower having the shape and location it has. We could not coherently conceive of the seven bricks being piled on top of one another in the way that they are in the absence of the tower. In contrast, it is eminently possible to conceive of our seven subjects of experience experiencing the colours of the spectrum, existing in the absence of a subject of experience having an experience of white. The existence of a group of spatial objects, O1 ... ON, with certain shapes and locations, can necessitate the existence of a spatial object with a shape and location different to the shape and location of each of O1 . . . ON. It does not seem that subjects of experience, merely in virtue of their existence, can stand in this kind of necessary relation.⁶

In the above passage I subscribe to the following epistemic principle:

Conceptual Isolation of Subjects (CIS)—For any group of subjects, instantiating certain conscious states, it is conceivable that just those subjects with those conscious states exist in the absence of any further subject.

From this I draw the following metaphysical conclusion:

Metaphysical Isolation of Subjects (MIS)—For any group of subjects, instantiating certain conscious states, it is possible that just those subjects with those states exist in the absence of any further subject.

In contrast, I rejected the following epistemic principle:

Conceptual Isolation of Lego (CIL)—For any group of Lego bricks, at certain locations, it is conceivable that just those objects at those locations exist in the absence of any further Lego object.

Hence I rejected the following metaphysical conclusion:

Metaphysical Isolation of Lego (MIL)—For any group of Lego bricks, at certain locations, it is possible that just those objects exist at those locations in the absence of any further Lego object.

I now think that in this earlier work I didn't get the disanalogy between the phenomenal case and the Lego case quite right. Whether or not James's nihilism about composite objects is plausible, it does not seem to be *incoherent*. One is not contradicting oneself when one claims that the Lego bricks do not really combine when arranged 'tower-wise' by a child, or that they combine only in the weak sense that they produce a distinct visual impression in the child playing with them. Perhaps such a view is counter to common sense, but it seems coherent.⁷ Thus, it now seems to me that *both* Conceptual Isolation of Subjects and Conceptual Isolation of Spatial Objects are true.

The crucial difference between Lego combination and subject combination arises when we try to move from conceivability to possibility. It would be helpful at this point to take a digression into the relationship between conceivability and possibility.

11.3 Conceivability and Possibility

Since Kripke moving from conceivability to possibility has not been so straightforward. In *Naming and Necessity* Saul Kripke argued that there are a

posteriori *necessities*, propositions which are necessarily true, but can only be known to be true through observation or experiment.⁸ The proposition water is H_2O is one example: we cannot know sitting in the armchair that water is H_2O , and yet there is no possible world in which water exists with some other chemical composition (in what follows I refer to concepts and propositions with underlined words). An a posteriori necessity is conceivably false, in the sense that we cannot know it to be true a priori, and yet is not possibly false. Thus, if there are a posteriori necessities, then a traditional principle of philosophy is false:

Conceivability Principle (CP)—If P is conceivably true, then P is possibly true.

The work of David Chalmers is perhaps best seen as an attempt to defend a somewhat traditional conception of philosophy in the light of Kripke's work. Chalmers accepts Kripke's examples of a posteriori identities and hence rejects CP. However, he defends a more subtle principle linking conceivability to possibility:

Two-Dimensional Conceivability Principle (2D-CP)—If P is conceivably true (upon ideal reflection), then there is a possible world W, such that P is true at W considered as actual.⁹

Twin Earth is the genuine possible world that is indiscernible from our world except that the colorless odorless stuff that fills oceans and lakes and falls from the sky is XYZ rather than H_2O . Chalmers agrees with Kripke that in our normal way of thinking about a possible world, considering it as *counterfactual*, that is, as a *way things might have been*, the proposition *water is XYZ* is false at Twin Earth. Water is the *actual* colorless, odorless stuff that fills oceans and lakes and falls from the sky; the actual stuff in oceans and lakes is H_2O ; and hence water is H_2O in all possible worlds (even ones where oceans and lakes and filled with XYZ).

However, Chalmers thinks that there is another way to think about a possible world: as *actual*, that is to say, as *a way things might actually turn out*. When Twin Earth is considered this way, the proposition <water is XYZ> comes out true at Twin Earth. Water is the actual stuff in oceans and lakes; if the actual stuff in oceans and lakes turns out to be XYZ, then water is XYZ. Hence, although necessarily false, there is a genuine possibility corresponding to <water is XYZ> in the sense articulated by 2D-CP: when that possibility is considered as actual, <water is XYZ> comes out true.¹⁰ Chalmers holds that every conceivably true proposition corresponds in this way to some genuine possibility.

I have argued in other work that 2D-CP rests on highly contentious semantic assumptions, for which Chalmers has not provided adequate defense.¹¹ However, I think that a much simpler and less contentious principle linking conceivability to possibility is both defensible and consistent with Kripke's cases: *conceivability entails possibility when you completely understand what you're conceiving of.* In the current context, we can partially explicate this principle in terms of the following:

Transparency Conceivability Principle (TCP)—For any proposition P, if (A) P involves only quantifiers, connectives, and predicates expressing transparent concepts, and (B) P is conceivably true (upon ideal reflection), then P is possibly true.

Understanding TCP requires understanding the distinction between *transparent* and *opaque* concepts, which I will try to express in the following.

For something to be water is for it to be constituted of H_2O molecules, but our ordinary concept of water does not reveal this. Our ordinary concept water is *opaque*, in the sense that it is not a priori (for someone possessing the concept, and in virtue of possessing the concept) what it is for something to be water. Due to the opacity of the concept <u>water</u>, one is able to conceive of scenarios involving water *without fully understanding what is being conceived of*. It is this lack of transparent understanding of what is being conceived of that blocks the move from conceivability to possibility.

In contrast to <u>water</u>, the concept <u>million-sided object</u> is *transparent*: it is a priori (for someone possessing the concept, and in virtue of possessing the concept) what it is for something to have a million sides. Hence when one conceives of a million-sided object one completely understands, or is in principle able to reason one's way to a complete understanding of, the situation being conceived of. In conceptions involving only quantifiers, connectives, and predicates expressing transparent concepts—for example a conception of there being a million-sided object—it is a priori for the conceiver what it is for the state of affairs they are conceiving of to obtain. In such cases, I believe that we can move from the conceivability (upon ideal reflection) of the states of affairs so conceived to its genuine possibility.

Suppose this were not the case. Suppose the existence of a million-sided object were just 'brutely impossible,' in the sense that, (A) a complete understanding of what it is for there to be a million-sided object, coupled with ideal reflection upon that state of affairs, cannot reveal any incoherence in there being a million-sided object; and, yet, (B) it is impossible for there to be a million-sided object. In order to illustrate this supposition we might imagine an omnipotent and perfectly ratio-nal being tries to create a million-sided object. She examines the notion of such a

thing from all angles and finds no bar to its existence. And yet when she tries to create such a thing, she finds herself unable. She is unable not because of any limit of power on her part but simply because it is impossible—in such a way that nobody could ever make intelligible sense of its being impossible—for there to exist a million-sided object.

When metaphysical possibility is so radically divorced from conceptual coherence, as vividly expressed in the last paragraph, I start to lose my grip on what metaphysical possibility is supposed to be. Moreover, a radical separation between what is conceivable and what is possible has the potential to make our knowledge of possibility problematic. If David Lewis is right that possibilities are concrete worlds spatiotemporally distinct from our own, then it is hard to account for our knowledge of what is or is not possible.¹² But if metaphysical possibility is just a special kind of conceivability, then knowing what is possible is just a matter of knowing what is conceivable in the requisite sense.

Furthermore, panpsychism is very often grounded in an opposition to physicalism, which is in turn very often grounded in some kind of conceivability argument. Hence, the panpsychist has strong motivational reason to accept some kind of principle linking conceivability to possibility. In defending panpsychism in the first section of this chapter, I was implicitly assuming the falsity of certain alternatives, such as physicalism about consciousness in conjunction with dispositional essentialism about physical entities. In a more extended defense of panpsychism, such alternatives would need to be ruled out.¹³

I take it, then, that the panpsychist has good reason to accept TCP, and I will assume it in what follows.¹⁴

11.4 Back to the Combination Problem

The predicate 'is a Lego brick' does not express a transparent concept. A Lego brick is essentially composed of a certain kind of plastic. That plastic is essentially constituted of a certain arrangement of atoms of certain kinds, and which atomic kinds and arrangements constitute the essence of plastic cannot be known a priori. Therefore, although I can conceive of Lego bricks being arranged 'tower-wise' in the absence of a Lego tower, in doing so I do not fully understand the nature of the state of affairs I am conceiving of. The real nature of a Lego brick is not apparent to me in my thoughts about Lego. On account of this, I cannot use TCP to move from the genuine conceivability of Lego bricks arranged tower-wise without a Lego tower (Conceptual Isolation of Lego) to the genuine possibility of Lego bricks arranged tower-wise without a Lego tower (Metaphysical Isolation of Lego).

In contrast, direct phenomenal concepts are plausibly transparent. One forms a direct phenomenal concept when one attends to a conscious state one is currently in and thinks about it in terms of how it feels.¹⁵ When I attend to a pain, it is directly revealed to me what it is for something to feel that way. When I attend to my experience of orange, it is directly revealed to me what it is for something to instantiate an experience of that kind. Many physicalists deny the thesis that direct phenomenal concepts are transparent (Phenomenal Transparency), holding that phenomenal concepts are entirely opaque concepts which turn out to denote brain states.¹⁶ However, on the assumption that phenomenal concepts are opaque, it is hard to make sense of the rich a priori knowledge which flows from phenomenal concepts, for example, the knowledge that phenomenal red is similar to phenomenal orange, or that pain has moral significance.¹⁷ Moreover, as discussed above, most panpsychists are motivated by an opposition to physicalism, commonly grounded in conceivability arguments. But if phenomenal concepts are opaque, then the kind of moves from consciousness-involving conceivability to possibility involved in the antiphysicalist arguments, facilitated by principles such as 2D-CP or TCP, are blocked.¹⁸ I will take it, then, that a panpsychist has good reason to accept phenomenal transparency.

Given phenomenal transparency, when I conceive of a group of subjects existing in the absence of a further subject, at least in so far as I have direct phenomenal concepts of the conscious states involved in my conception, I completely understand what I am conceiving of. It follows from phenomenal transparency that the conception we reach when we reflect on the famous James passage involves (or at least can involve) only quantifiers, connectives, and predicates expressing transparent concepts.

Thus, we reach the heart of the combination problem: given the lack of opacity in the relevant conception, we are licensed to infer from Conceptual Isolation of Subjects to Metaphysical Isolation of Subjects. We may press the difficulty with the following argument:

The *No Summing of Subjects Argument* (the heart of the combination problem)

- 1. *Conceptual Isolation of Subjects*—For any group of subjects, instantiating certain conscious states, it is conceivable that just those subjects with those conscious states exist in the absence of any further subject.
- 2. *Transparency Conceivability Principle*—For any proposition P, if (A) P involves only quantifiers, connectives, and predicates expressing transparent concepts, and (B) P is conceivably true upon ideal reflection, then P is metaphysically possibly true.

- 3. *Phenomenal transparency*—Phenomenal concepts are transparent.
- 4. *Metaphysical Isolation of Subjects*—For any group of subjects, instantiating certain conscious states, it is possible that just those subjects with those states exist in the absence of any further subject (from 1, 2, and 3).
- 5. For any group of subjects, those subjects with those conscious states cannot account for the existence of a further subject (from 4).
- 6. Therefore, panpsychism is false (from 5).

11.5 Phenomenal Bonding—A Response to the No Summing of Subjects Argument

I believe the No Summing of Subjects Argument to be sound right up to 4, which is to say that there is a sound argument for Metaphysical Isolation of Subjects. But let us think more carefully about what implications MIS has for the summing of experiences. It follows from MIS that certain subjects of experience cannot sum merely in virtue of their existing (and instantiating the specific phenomenal characters they instantiate). But it does not imply that a certain set of subjects of experience cannot exist and be involved in some *state of affairs* which accounts for the existence of some distinct subject of experience. There is nothing in the principle that rules out the possibility of there being some state of affairs of certain subjects of experience *being related in some specific way* which necessitates the existence of some distinct subject of experience.

To put it another way, MIS implies that there is no state of affairs of the form <subject of experience S1 exists with phenomenal character x, and subject of experience S2 exists with phenomenal character y> which necessitates <subject of experience S3 exists with phenomenal character z>. But it does not imply that there is not some state of affairs of the form <subject of experience S1 with phenomenal character x bears relationship R to subject of experience S2 with phenomenal character y> which necessitates <subject of experience S2 with phenomenal character y> which necessitates <subject of experience S3 exists with phenomenal character z>. Such a sense of experience summing is not ruled out by MIS.¹⁹

I don't think we have a transparent conception of such a relation, call it 'phenomenal bonding,' which bonds subjects together to produce other subjects of experience. If we did have such a conception, then the solution to the combination problem would be obvious. Indeed, the problem would never have occurred to us.

However, it is not surprising that we lack a transparent grasp of the phenomenal bonding relation—if such a thing there be—given the nature of our epistemic situation. Our most basic empirical science, physics, yields understanding only of the world's mathematico-causal structure, and the phenomenal bonding relation is not a mathematico-causal relation: conceiving of subjects standing in mathematico-causal relations does not remove their conceptual isolation, and hence does not remove their metaphysical isolation. Apart from its mathematico-causal structure, arguably the only feature of the world we transparently understand is consciousness. And consciousness is a monadic property. Our unfortunate epistemic situation does not afford us a transparent understanding of the (non-mathematico-causal) relations which conscious things bear to each other.

Here is another way of pressing the point. Human beings are able to have neither introspective nor perceptive experience of relations between subjects of experiences qua subjects of experience. We are unable to perceive relations between subjects of experience (qua subjects of experience) through the senses simply because we are unable to perceive subjects of experience (qua subjects of experience) through the senses. If you examine my brain, you will not be able to see it instantiating phenomenal properties. I have epistemic access to only one subject of experience qua subject of experience, that is, the subject of my own experience accessed via introspection. And it follows from the fact that we can introspect only one subject of experience that we cannot introspect how subjects of experience qua subjects of experience are related, for to introspect how subjects of experience qua subjects of experience are related we would have to be able to introspect more than one subject of experience. Given that we can experience subjects of experience qua subjects of experience only via introspection, and we have introspective access only to one subject of experience, it follows that we cannot experience subjects of experience qua subjects of experience as related.

Just because we are unable to form a transparent conception of the phenomenal bonding relation does not mean we cannot form a conception of it. We can think of it as 'the relation such that when subjects stand in it they produce a further subject,' and we can suppose that there is such a thing. We may even be able to identify it with some relation we can observe in the world or some relation that features in physics. None of the relations that appear in perception or in physics is conceived of *as* a phenomenal bonding relation. In the same way, the brain does not appear from the outside *as* a subject of experience, and the properties of physics or neuroscience are not conceived of in those sciences *as* phenomenal qualities. But just as the panpsychist might identify charge with a form of consciousness, so the proponent of phenomenal bonding might identify some empirically known relation as the phenomenal bonding relation.

I can see no principled reason to think the phenomenal bonding relation is not a real relation that certain subjects bear to each other, and I think therefore we have a way of making sense of subjects summing, and hence a way of making sense of panpsychism. The theoretical attractions of panpsychism give us good reason to take this route to saving the view and hence to believe that there is a phenomenal bonding relation.

The flaw of the No Summing of Subjects Argument, then, is in the final stages. We can only move from Metaphysical Isolation of Subjects to

5. For any group of subjects, instantiating certain conscious states, those subjects with those conscious states cannot account for the existence of a further subject.

if we construe 'account for the existence of' as meaning 'wholly explain the existence of just by their mere existence and intrinsic conscious nature.' And we can only infer the falsity of panpsychism from this construal of 5 if we construe panpsychism as requiring that subjects 'wholly explain just by their mere existence and intrinsic conscious nature' the existence and nature of other subjects. But there is no reason to construe panpsychism in this way. The nature of organisms and car engines are accounted for in terms of their parts, but those parts constitute the organism/engine only when related in the right way. The same is surely true of the explicability of subjects in terms of other subjects.

There is a sense in which embracing this solution to the combination problem leads to a kind of mysterianism. In so far as we don't have a transparent grasp of the phenomenal bonding relation, there is a clear sense in which we don't understand how subjects combine. This lack of knowledge is frustrating, which may cause us to yearn for a different theory. However, as hard-nosed metaphysicians, we should be asking not which view we'd like to be true, but which view is most likely to be true. And the great elegance with which panpsychism unifies the existence of consciousness with the facts of observation renders it highly likely to be true. Probably the success of physics, and the collective forgetting that all physics gives us is structure, has given us rather high expectations of what we can achieve concerning our understanding of the universe. But once we get our epistemic situation right in the way I described earlier, we are more humble in our aspirations.

Indeed, independently of the desire to make sense of panpsychism, we have good reason to think that there are relations we lack a transparent conception of. We know that things stand in spatiotemporal relations, and yet physics provides us only with a mathematical conception of such relations. Mathematical descriptions abstract from the concrete nature of things. We abstract from the concrete reality of phenomenology when we describe it merely in terms of its mathematical structure, for example, we abstract from the real concrete nature of five subjects when we describe them as merely 'five things.' Intuitively, there must be some real nature to spatiotemporal relations underlying the mathematical understanding of those relations we get from physics. If that intuition is correct, then we are completely in the dark about the real nature of spatiotemporal relationships.

11.6 Intelligible Emergentist Panpsychism Versus Constitutive Panpsychism

The kind of consciousness we want a theory of consciousness to explain is the kind of consciousness we pretheoretically associate with humans and other animals. I call this kind of consciousness 'o-consciousness' ('o' for 'ordinary' or 'organism'). I have tried in the above to make sense of the idea that o-conscious subjects are intelligibly produced by the consciousness and the phenomenal bonding of microlevel entities. However, I do not take intelligible production to be the same thing as constitution. I take cases of constitution to be a subset of cases of intelligible production; the defining characteristic of constitution being that constituted states of affairs are *nothing over and above* the states of affairs which constitute them.

Consider the following examples of intelligible production that do *not* involve constitution. We can move a priori from *God willing that there be light* to *there being light*, and the latter state of affairs obtains because the former obtains; in this sense the state of affairs of there being light is intelligibly produced by the state of affairs of God willing that there be light. Nonetheless, there being light is not constituted of God's willing that there be light; in willing that there be light God creates new being. Similarly, if the conjunction of dispositional essentialism and determinism is true, we can move a priori from facts about the past to facts about the future, but it does not follow that facts about the future are constituted of facts about the past.

Thus, we can distinguish two forms of panpsychism, both of which involve facts about o-consciousness being intelligibly produced by consciousness-involving facts at the micro-level:

Constitutive panpsychism—Facts about o-consciousness are constituted by, and hence are nothing over above, consciousness-involving micro-level facts.

Intelligible emergentist panpsychism—Facts about o-consciousness are intelligibly produced by, but are something over and above, consciousness-involving micro-level facts.

David Chalmers has argued that constitutive forms of panpsychism enjoy an important advantage over nonconstitutive forms, in so far as they are able to reconcile the causal efficacy of o-consciousness with the causal closure of the microphysical.²⁰ In "Against Constitutive Russellian Monism," I argue against the constitutive view, whilst in "Consciousness and Fundamental Reality" I argue

that only cosmopsychist forms of the constitutive view are viable.²¹ In any case, it is important to note that the case I made for panpsychism in the first section of this paper, grounded in considerations of theoretical elegance, is neutral between constitutive and emergentist forms of panpsychism.

11.7 When Do Subjects Sum?

Peter van Inwagen encouraged metaphysicians to ask 'the special composition question': Under what conditions do objects combine to form a further object?²² Panpsychists can ask 'the special *phenomenal* composition question': Under what conditions do subjects combine to produce a further subject? For believers in phenomenal bonding, this will be equivalent to the question: Which subjects bear the phenomenal bonding relation to each other?

One popular answer to the special composition question is 'always.²³ For proponents of *unrestricted composition* any objects, no matter how disparate and seemingly unrelated, compose an object: your nose, my teeth, and the planet Venus form an object. Other philosophers defend some form of *restricted composition*: some sets of objects are such that their members compose an object, other sets of objects are such that their members do not compose an object.²⁴ We can distinguish analogous answers with respect to the special phenomenal composition question. According to unrestricted phenomenal composition, for any group of subjects, say, the particles forming your nose, my teeth, and the planet Venus, those subjects are related by the phenomenal bonding relation and hence produce a further subject. Obviously, some form of restricted phenomenal composition, according to which some but not all subjects are such that they bear the phenomenal bonding relation to each other, will be more in keeping with pre-theoretical common sense.

One's approach to answering the special phenomenal composition question will vary radically depending on whether one is a constitutive or an emergentist panpsychist. If emergentism is true, and if emergent entities have distinctive causal powers, then there will be an empirically discernible distinction between those systems that have and those systems that lack such emergent causal powers. The behavior of the latter but not the former will be predictable on the basis of the behavior of the system's parts. Hunting the phenomenal bonding relation, for the emergentist, will be a matter of looking for an empirically distinguished relation that relates the parts of systems with emergent causal powers but doesn't relate the parts of systems that lack emergent causal powers. It is likely, then, that the emergentist will support some form of restricted phenomenal composition, looking to the empirical facts for the boundary between systems that are mere aggregates of micro-level subjects and systems that are conscious in their own right. For the constitutive panpsychist, in contrast, higher-level subjects are nothing over and above microsubjects, and their causal powers are entirely derived from the causal powers of their parts. It is less clear in this case that there will be an empirically discernible distinction between systems which have and systems which lack consciousness in their own right. We cannot directly observe either the presence or absence of consciousness in a system. Sciences such as psychophysics are reliant on the anti-skeptical assumption that what a subject says about her consciousness is a reliable guide to the facts about her consciousness. Such anti-skeptical assumptions give us reason to believe in the presence of consciousness in certain material systems, for example systems that talk to us, but they give us no reason to believe in the *absence* of consciousness in any material systems. To avoid skeptical scenarios we must attribute consciousness to organic systems, but we are not required to refrain from attributing consciousness to non-organic systems.

What can we turn to if observation can't help us answer the special phenomenal composition question? It is probably the case that most people researching consciousness draw their answer to the special phenomenal composition question from pre-theoretical common sense, taking it to be something like the following:

Commonsense answer—Particles form a conscious subject when and only when they form an organism (or a subset of organisms, or the brains/central nervous systems of organisms; I will ignore these subtleties for the sake of simplicity).

Unfortunately, there is a difficulty, arising from considerations of vagueness, with accepting this answer. In what follows I will outline this difficulty.

The boundary between the organic and the nonorganic is vague. There are what we can call 'organic borderline cases'—cases where there is no fact of the matter as to whether or not we have a human organism—at the beginning and end of an organism's existence. In any particular case there is no utterly precise point in time, after which we have a zygote, and before which we have only sperm and egg. Similarly, in each particular case there is no utterly precise point in time after which we have a corpse and before which we have a living body. Given our macroscopic concerns, this vague boundary is barely discernible. But if we were looking at a complete description of the fundamental particles composing a human organism during, and slightly before and after, its existence, there would be no precise arrangement of particles which constituted the beginning and end of the organism's existence; there would be borderline cases. There would presumably be similar borderlines cases at the coming to be/passing away of brains, or coming to be/passing away of brains of the level of complexity allegedly required for macro-level consciousness.

If the commonsense answer to the special phenomenal composition question is correct, it follows that there are 'phenomenal borderline cases,' cases where there is no fact of the matter whether we have a conscious subject. If the existence of an organism is necessary and sufficient for the existence of a conscious subject, and if it's sometimes vague whether we have an organism, it follows that it's sometimes vague whether we have a conscious subject.

Why should this be thought to be a problem? Why should a vague boundary between the conscious and the nonconscious be any more problematic than a vague boundary between the tall and the non-tall, or the heap and the nonheap? The trouble arises if we want to give a semantic treatment of vagueness. Making sense of vagueness as a semantic phenomenon requires associating each vague predicate with a spectrum of sharpenings, and (given a couple of fairly plausible assumptions) it is hard to make sense of the predicate 'is conscious' being associated with a spectrum of sharpenings.²⁵

Let us take this more slowly. According to semantic theories of vagueness, vagueness is the result of *semantic indecision*: for any vague predicate there are multiple 'sharpenings' of the predicate, such that the meaning of the predicate does not settle on any of these sharpenings. Consider the vague predicate 'is tall.' We could stipulate, somewhat arbitrarily, that anything that is exactly six feet or taller counts as 'tall,' and anything shorter is not tall. This is one 'sharpening' of the predicate 'is tall,' that is, one way of making the predicate precise. Alternately, we could stipulate that anything that is exactly six feet and one inch or taller counts as tall, and anything shorter is not tall. This is an alternative sharpening of 'is tall,' that is, an alternative possible way of making the predicate precise. The predicate 'is tall' is thus associated with a *spectrum of sharpenings*: a range of possible ways of making the predicate precise.²⁶

Semantic theories of vagueness tell us that a vague predicate is vague because no one has bothered to single out one of its sharpenings as the unique meaning of the predicate. To put it metaphorically, the predicate hasn't made up its mind which of those precise meanings it wants to plump for. Suppose John is a borderline case of tallness. According to semantic theories of vagueness, it's not that in reality there is some fuzzy, indeterminate state of affairs of John's neither having nor lacking a certain quality. In the world there's just John with some utterly precise height. It's *the predicate* that is indeterminate such that there's no fact of the matter as to whether it applies to things with John's exact height. The indeterminacy is in language rather than the world.

There seems to be two ways of making sense of the predicate 'is conscious' being associated with a spectrum of sharpenings. The first is to adopt analytic functionalism, according to which it is a priori that for something to be conscious is for it to instantiate a certain functional state. On such views, it is natural to take the predicate 'is conscious' to be indeterminate with respect to a range of slightly different functional states, each of which could be captured with a fine-grained enough description. The second strategy is to hold that the meaning of the predicate 'is conscious' is determined by facts outside of what is a priori accessible to the concept user; and so, although the meaning of the predicate involves a spectrum of sharpenings, that spectrum of sharpenings is not a priori accessible.²⁷

The panpsychist is unlikely to be able to adopt either of these strategies. The first strategy entails a deflationary account of consciousness, whilst panpsychism is usually grounded in a robust metaphysical seriousness about consciousness. The second strategy involves the rejection of phenomenal transparency, and as I have argued above the panpsychist has good reason to accept *phenomenal transparency*.

If we are unable to give a semantic treatment of vagueness, then adopting the common sense answer to the special phenomenal composition question will require thinking of phenomenal borderline cases in metaphysical or epistemic terms. I haven't the courage to wade too much into the overwhelming literature on vagueness, but having dipped in my toe, I find myself strongly inclined toward a semantic account of vagueness,²⁸ and thereby inclined against the commonsense answer to the special phenomenal composition question. Others must weigh these considerations for themselves.²⁹

There is a strong possibility, then, that neither the facts of observation, nor the facts of common sense will be able to help the constitutive panpsychist answer the special phenomenal composition question. In such a case, she must turn to theoretical virtue. On account of their elegant simplicity, the two most theoretically satisfying answers to the special phenomenal composition question are *nihilism*—subjects never combine to make a further subjects, and universalism (unrestricted phenomenal composition)—subjects always combine to make a further subject. Therefore, if we can take it that nihilism is a nonstarter on the grounds that the subjects we are pre-theoretically committed to are composite objects of some sort, universalism looks to be the default position. In a similar way, before we had empirical evidence to the contrary, the default assumption was that the speed of light was infinite (zero and infinity being the simplest values). Arbitrary constants are to be avoided if at all possible. Universalism is wildly at odds with common sense, but we have shown that there are deep problems with what common sense has to tell us about phenomenal composition.

If we are prepared to accept universalism, there is an obvious candidate for identification with the phenomenal bonding relation: the spatial relation. If we identify the phenomenal bonding relation with the spatial relation it follows that, for any group of material objects, the members of that group, being spatially related, constitute a conscious subject. A nice consequence of identifying the phenomenal bonding relation with the spatial relation is that we end up having some positive conception of the spatial relation. As noted at the end of the section 5, the spatial relation must have some real nature that goes beyond the mathematical conception of it we get from physics. Supposing that the real nature of the spatial relation is the phenomenal bonding relation is not entirely satisfying, as we lack a transparent conception of that relation: we don't completely understand what it is for objects to be related in that way. Still, it is better than nothing; at least we know that the real nature of the spatial relation is such as to bond subjects together to constitute further subjects. On the view currently under consideration, we have a reasonable grip on the nature of the world: the only intrinsic determinable is consciousness, the only relational determinable is phenomenal bonding.

11.8 Conclusion

Some form of panpsychism is highly likely to be the true theory of our universe. It's high time we started working out the details.³⁰

Notes

- Of course there are a number of options I am implicitly ruling out here, such as physicalism about consciousness in conjunction with dispositional essentialism about physical entities. I argue against these views elsewhere (Goff 2011a; 2015a; 2015b; forthcoming). Here I would like to concentrate on exploring the details of panpsychism.
- 2. Chalmers (this volume) spells out a whole series of combination problems, grouped around the three themes of subjective character, qualitative character, and structural character. I think the central problem pertaining to subjective character, that is, the difficulty I explore here of making sense of subjects summing, is recognized to be the central 'combination problem,' and the one that threatens to undermine the coherence of panpsychism from the off. The problems pertaining to qualitative and structural character strike me more as challenges that need to be responded to when working out the details of a specific panpsychist view (as Coleman does in this volume), rather than potentially yielding knock down arguments against panpsychism.
- 3. James 1890, 1:160.
- 4. James 1890, 1:158–59.
- 5. For a good discussion of James' argument, see Shani 2010.
- 6. Goff 2009a, 130-31. I further defend CIS (although not under that name) using panpsychist zombies in Goff (2009b). Chalmers discusses this argument in his chapter in this volume.
- 7. Theodore Sider (2013) has recently defended the non-existence of composite objects.
- 8. See Kripke 1972/1980.
- 9. See Chalmers 1996; 2002; 2009.
- 10. Ultimately Chalmers grounds all this in what it is rational to suppose. On the assumption that the stuff in oceans and lakes is XYZ, it is rational to suppose that water is XYZ.
- 11. See Goff and Papineau 2014; Goff 2011.

- 12. See Lewis 1986.
- 13. I try to rule out these alternatives in Goff 2015a; 2015b; forthcoming..
- 14. I provide further defense of a less-restricted version of TCP—able to deal with propositions involving singular terms—in Goff and Papineau (2014) and Goff (*forthcoming*).
- 15. The notion of a direct phenomenal concept is from Chalmers (2003).
- 16. McLaughlin (2001) and Papineau (2006) explicitly deny phenomenal transparency. Most contemporary physicalists adopt some kind of semantic externalist account of the reference of phenomenal concepts, which seems to entail the falsity of phenomenal transparency: if reference to a given conscious state is determined by facts external to our understanding, then it's hard to see how we could have a priori access to the essence of consciousness.
- 17. I argue at length against physicalist views of phenomenal concepts, on the basis of these kinds of consideration, in Goff (2015a; forthcoming).
- 18. In relation to 2D-CP, moves from conceivability involve a concept having identical primary and secondary intensions (i.e., having the same referent in a world whether that world is conceived of as actual or as counterfactual). Given that the primary intension can be evaluated a priori, and the secondary intension reflects the essence of the referent, this is more or less equivalent to the concept's being transparent.
- 19. Goff(2009a) suggests this solution to the combination problem.
- 20. See Chalmers this volume; 2015.
- 21. See Goff 2015b, forthcoming. The argument I present in this chapter might be seen as a way of pressing a version of the combination problem against constitutive Russellian monism. Interpreted in this way, as a challenge to constitutive panpsychism, I take the combination problem to be insoluble. Interpreted as a challenge to the more general view that higher-level phenomenal facts can be intelligibly determined by more basic phenomenal facts, I take the combination problem to be soluble.
- 22. See van Inwagen 1990.
- 23. See Lewis 1986; Sider 2001.
- 24. See van Inwagen 1990; Merricks 2001.
- 25. See Dummett 1978, 260; Fine 1975; Russell 1923; Lewis 1986, 212.
- 26. With some vague predicates, as with 'is tall,' the sharpenings are determinates of a single determinable. In the case of other vague predicates, e.g., 'is a religion,' there is a weighted cluster of properties, *involves belief in a supernatural being, involves ritual, involves a moral code*, such that each sharpening involves some of those properties, but it is not the case that each sharpening involves all of those properties.
- 27. See Papineau 2002.
- 28. If find persuasive Horgan and Potrc's (2008) argument against the coherence of metaphysical vagueness. For a contemporary account of metaphysical vagueness, see Barnes and Williams (2011).
- My formulation of these considerations is influenced by Lewis's (1986) and Sider's (2001) arguments against unrestricted composition and is similar to arguments I gave in Goff (2011b) and Goff (2013).
- 30. I am grateful to Luke Roelofs, Hedda Hassel Mørch, and David Chalmers for comments.

Bibliography

- Barnes, Elizabeth, and J. Robert G. Williams. "A Theory of Metaphysical Indeterminacy." Oxford Studies in Metaphysics. Ed. Karen Bennett and Dean W. Zimmerman. Oxford: Oxford University Press, 2011. 103–48.
- Chalmers, David. The Conscious Mind: Towards a Fundamental Theory. Oxford: Oxford University Press, 1996.
- Chalmers, David. "Consciousness and Its Place in Nature." Philosophy of Mind: Classical and Contemporary Readings. Ed. David Chalmers. Oxford, New York: Oxford University Press, 2002. 247–72.

- Chalmers, David. "The Content and Epistemology of Phenomenal Belief." *Consciousness: New Philosophical Perspectives.* Ed. Qentin Smith and Aleksandar Jokic. Oxford: Oxford University Press, 2003. 220–72.
- Chalmers, David. "The Two-Dimensional Argument Against Materialism." Oxford Handbook of Philosophy of Mind. Ed. Ansgar Beckermann, Brian P. McLaughlin, and Sven Walter. Oxford: Oxford University Press, 2009. 313–39.
- Chalmers, David. "The Combination Problem for Panpsychism." *Panpsychism*. Ed. Godehard Brüntrup and Ludwig Jaskolla. Oxford: Oxford University Press, this volume.
- Chalmers, David. "Panpsychism and Panprotopsychism." *Russellian Monism*. Ed. Torin Alter and Yujin Nagasawa. Oxford: Oxford University Press, 2015.
- Dummett, Michael. Truth and Other Enigmas. London: Duckworth, 1978.
- Fine, Kit. "Vagueness, Truth and Logic." Synthese 30 (1975): 265-300.
- Goff, Philip. "Can the Panpsychist Get Round the Combination Problem?" Mind that Abides: Panpsychism in the New Millennium. Ed. David Skrbina. Amsterdam: Benjamins, 2009a. 129–136.
- Goff, Philip. "Why Pansychism Doesn't Help Explain Consciousness." *Dialectica* 63 (2009b): 289–311.
- Goff, Philip. "A Posteriori Physicalists Get our Phenomenal Concepts Wrong." Australasian Journal of Philosophy 89 (2011a): 191–209.
- Goff, Philip. "There Is No Combination Problem." *The Mental as Fundamental*. Ed. Michael Blamauer. Heusenstamm: Ontos, 2011b. 131–40.
- Goff, Philip. "Orthodox Property Dualism + Linguistic Theory of Vagueness = Panpsychism." Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience. Ed. Richard Brown. Heidelberg: Springer, 2013. 75–92.
- Goff, Philip. "Real Acquaintance and the Phenomenal Concepts Strategy." *The Nature of Phenomenal Qualities*. Ed. Paul Coates and Sam Coleman. Oxford University Press, 2015a.
- Goff, Philip. "Against Constitutive Russellian Monism." *Russellian Monism*. Ed. Torin Alter and Yujin Nagasawa. Oxford: Oxford University Press, 2015b.
- Goff, Philip. Consciousness and Fundamental Reality. New York: Oxford University Press, forthcoming.
- Goff, Philip, and David Papineau. "What's Wrong with Strong Necessities?" *Philosophical Studies* 167 (2014): 749–62.
- James, William. The Principles of Psychology. 2 vols. Mineola, NY: Dover, 1890/1950.
- Kripke, Saul. "Naming and Necessity." Semantics of Natural Language. Ed. David Davidson and Gilbert Harman. Dordrecht: D. Reidel, 1972. Rpt. as Naming and Necessity. Cambridge: Harvard University Press, 1980.
- Lewis, David. On the Plurality of Worlds. Oxford: Blackwell, 1986.
- McLaughlin, Brian P. "A Defense of New Wave Materialism: A Response to Horgan and Tienson." *Physicalism and Its Discontents*. Ed. Carl Gillett and Barry Loewer. Cambridge: Cambridge University Press, 2001. 319–30.
- Merricks, Trenton. Objects and Persons. Oxford: Oxford University Press, 2001.
- Russell, Bertrand. "Vagueness." Australasian Journal of Philosophy and Psychology 1 (1923): 84-92.
- Sider, Theodore. Four-Dimensionalism: An Ontology of Persistence and Time. Oxford: Oxford University Press, 2001.
- Sider, Theodore. "Against Parthood." *Oxford Studies in Metaphysics*. Ed. Karen Bennett and Dean W. Zimmerman. Vol. 8. Oxford: Oxford University Press, 2013.
- Shani, Itay. "Mind Stuffed with Red Herrings: Why William James' Critique of the Mindstufftheory does not Substantiate a Combination Problem for Panpsychism." Acta Analytica 25 (2010) (4): 413–34.
- Papineau, David. Thinking About Consciousness. Oxford: Oxford University Press, 2002.
- Papineau, David. "Comments on Galen Strawson." Journal of Consciousness Studies 13 (2006): 100–109.
- Van Inwagen, Peter. Material Beings. Ithaca: Cornell University Press, 1990.

PART IV

PANPSYCHISM AND ITS ALTERNATIVES

12

Mind Dust, Magic, or a Conceptual Gap Only?

BRIAN P. MCLAUGHLIN

After the transgression in the Garden that sealed our fate, God reminded Adam: "Dust thou art, and unto dust thou shalt return" (Genesis 3:19). Modern science has confirmed that we are indeed wholly composed of dust—of atoms, dust from the stars.¹

We also each know, or at least should know in our hearts, that we are sentient, that we are phenomenally conscious. Material bodies experience acceleration and the like. But we know from our own case that it is like something for us to have certain experiences. We have subjective experiences, experiences that are like something for us as subjects. The experiences of feeling pain and of having a visual experience are paradigm examples. It is thus like something to be us. A being is phenomenally conscious just in case it is like something to be that being (see Sprigge 1971; Nagel 1974/1979). It is like something to be a being just in case the being has subjective experiences. We are phenomenally conscious, and so subjects of experience. So are the beasts of the Earth. There is, for instance, something that it is like to be a bat. Whether the creepy things that creep the Earth are phenomenally conscious remains open. Sense perception may not require sense experiences. Butterflies, for instance, see. They even have color constancy, and so, if cognitive science can be trusted, visual representations. But it's another matter whether they have visual experiences. It may well be that there's nothing that it is like to be a butterfly. (Cognitive science tells us that we have visual representations in our dorsal visual system, but what happens there is not associated with visual experience (Goodale and Milner 1992; 2005).) Some animals, though, including dolphins, elephants, and apes, are, like us, not just phenomenally conscious but also self-conscious. But of the beings we've thus far encountered—the material beings at any rate—we alone possess language and the ability to reflect on our place in nature.

We can see by reflection that we face a problem when we try to locate ourselves in nature. How could dust form a phenomenally conscious being? We can distinguish aggregates of dust particles from mereological sums of them and systems of dust from mere aggregates. We and the beasts of the Earth are organisms, and so systems of dust particles. But how could even a system of dust particles be phenomenally conscious? How could a phenomenally conscious being, a subject of experience, just consist in a system of dust? How could an object, no matter how complexly organized, be a subject? How could an *en-soi* be a *pour-soi*? Could it really be that we subjects of experience are dust in the wind of the forces of nature?

One answer is that Genesis and modern science are wrong. We aren't wholly composed of dust. We have an immaterial part—a soul or mind. That is the position taken by Augustine under the influence of Plato's philosophy (van Inwagen 1995). But if soul and body, or mind and brain are separate, what unites them? Augustine answers: "The manner in which spirits are united to bodies is altogether wonderful and transcends the understanding of men" (cited in Haldane 1994, 335). In modern lingo: we're *cognitively closed* to the properties that explain the nexus, in the way that a bat, say, is cognitively closed to the property of being a differential equation. It's a matter beyond our ken (McGinn 1989). Not everyone is convinced. (When is everyone convinced?) Much of the history of the modern mind-body problem from Descartes on has been an attempt to determine how immaterial minds are related to brains. But there remains to this day no satisfactory answer.²

I'll here put my faith in *Genesis* and simply assume we are wholly composed of dust, and so devoid of an immaterial soul or mind. But how then could a phenomenally conscious being, a subject of experience, just consist in a system of dust? How could it be that dust is so propertied and organized that there is a subject of experience?

One answer is that the particles of dust are themselves phenomenally conscious, and so phenomenal consciousness is ontologically fundamental, rather than constituted by something else. The dust of which we are composed is physical but also endowed with phenomenal consciousness. It is mind dust: physical dust with a conscious mind. The bits of star dust that make us up are themselves subjects of experience, not merely objects. It's like something to be them. They have subjectivity. Subjectivity is present at the basic ontological level rather than constituted somehow by something wholly objective.

Although this view doesn't entail that literally everything (numbers included), or even everything in spacetime (and so, galaxies, dark matter, the mereological sum of the Big Bang and the Big Crunch), is phenomenally consciousness, it still plausibly counts as a kind of panexperientialism, and so a kind of panpsychism, a doctrine that has a venerable philosophical history (see Skrbina 2005). Given that a being is phenomenally conscious just in case it is like something for that being as a subject to have certain experiences, and so certain of the experiences of the being are subjective experiences, this kind of panexperientialism entails a kind of pansubjectivism, even though not literally everything falls into the pan.

I'll focus on an argument of Thomas Nagel for what he takes to be a kind of panpsychism, namely, "the view that the basic physical constituents of the universe have mental properties" (Nagel 1979, 181). It's my focus, because it's the best argument that I've seen for that view.³ The argument is flawed. But the flaws are relatively minor. Once removed, only very difficult, substantive issues remain, issues concerning the truth of two of the premises, both of which, at first blush at least, seem plausible.

Although Nagel frames the argument in terms of mental properties, I'll recast it by restricting it specifically to properties of phenomenal consciousness, and so will have panexperientialism, rather than panpsychism, as its conclusion. By properties of phenomenal consciousness, I mean qualia in one of the (many) philosophical uses of that term. Qualia, in the intended sense here, are the what-it-is-like for the subject aspects of subjective experiences. There must be such, given that subjective experiences are like something for the subject having or undergoing them. Indeed, an experience is subjective if and only if it has qualia; if and only if it is like something for the subject of the experience. We can type experiences by their qualia; thus qualia are types of experiences. A quale such as the feel of pain is a type subjective experience. This recasting of the argument is not a departure from Nagel's intended argument for panpsychism. He makes it clear that the argument doesn't apply to mental states like beliefs (Nagel 1979, 182); that it concerns "conscious aspects" of mental states (Nagel 1979, 183), mental states that imply "the consciousness of [their] subject" (Nagel 1979, 183), "the subjective features" of mental states (Nagel 1979, 188), "the phenomenological qualities of our experiences" (Nagel 1979, 188), "what any conscious mental state is like for its possessor" (Nagel 1979, 188). I take him to mean qualia in our sense: an experience's being like something for a subject, and so being a subjective experience. (There are other notions of qualia in the literature.)

I'll first present Nagel's argument so recast (and with some additional minor changes), and then make some revisions to repair flaws. Once the revised argument is in place, we'll see that if panexperientialism, pansubjectivism, is mistaken, then we seem to face a stark choice, one between two alternatives, both of which themselves face truly formidable difficulties. I'll then examine the alternatives.

Here is Nagel's argument recast as an argument for panexperientialism (Nagel 1979, 181-82):

- **Premise One**. Human beings are complex systems wholly composed of physical particles.
- **Premise Two**. Qualia are not logically implied by physical properties alone. **Premise Three**. Human beings have qualia.
- **Premise Four**. There are no emergent properties. That is to say, all properties of a complex system that are not relations between it and something else derive from the properties of its constituents and their effects on each other when so combined.

Conclusion. The basic physical constituents of the universe have qualia.

I'll take premise one to claim that human beings are complex systems wholly composed of physical particles in the sense that they are constituted at any time (and throughout any interval of time) by some or other complex system so composed.⁴ The premise is thus consistent with the claim that human beings and such systems have different persistence conditions. Human beings are, of course, *homo sapiens*, the only extant members of the *hominin clade*. A physical duplicate of a complex system constituting a human being could fail to constitute a human being. If wet-life artificial life researchers someday synthesize a sperm and egg just from chemical elements, and combine them to grow a physical duplicate of a human being in a lab, the lab grown individual would not be a human being. That is compatible with premise one. Premise one entails that every human being (at a time or throughout an interval of time) is constituted by some system of physical particles, not that being a human being is identical with being a certain kind of system of physical particles.

'Wholly composed of' in premise one doesn't mean 'only composed of.' Premise one is compatible with the fact that we are organisms wholly composed of cells. Premise one requires only that we are wholly composed of physical particles.

I'll sometimes speak of levels of decomposition. By levels I just have in mind (roughly delineated) spatiotemporal scales, so that the constituents of something at a certain level are things at that scale that compose it. Something is wholly decomposable at a level, decomposable into things at a certain scale, just in case there are nonoverlapping entities at that scale that wholly compose it. On the evidence, human beings are wholly decomposable into entities at more than one scale. We are wholly composed of cells that are wholly composed of molecules that are wholly composed of atoms, which are wholly composed of electrons, protons, and (save hydrogen) neutrons, and, in turn, neutrons and protons are wholly composed of quarks, which are

themselves fundamental particles or else wholly composed of more fundamental physical entities. The relation of being a constituent or part of something is not, in general, transitive. I'm part of the club. The atom, 'Sparky,' is part of me. Although Sparky is attending the club meeting with me, Sparky isn't part of the club. Still, it may be that the material parthood relation is transitive. In any case, premise one implies that we are wholly decomposable into physical particles, and so wholly decomposable into physical entities at the smallest scale.

It may of course turn out that fundamental physical entities aren't particles, but instead vibrating strings of energy, or branes, or something else. I'll deal with that later. For now, let's just proceed with the assumption that there are physical particles.

Premise two states that qualia are not logically implied by physical properties alone. Logical entailment is standardly understood as a relationship between propositions, or statements, or claims, or sentences in a context of use. The notion of property entailment can be understood, however, as follows: A property Φ logically implies a property Ψ if and only if it is logically necessary that whatever has Φ has Ψ .⁵

It should be apparent that in the relevant sense of logical necessity, a truth can be logically necessary without being a logical truth. The notion is not a formal one. It's a metaphysical and epistemological notion. It is logically necessary that whatever has Φ has Ψ , if and only if (a) it is metaphysically necessary that (true in literally every possible world that) whatever has Φ has Ψ , and, (b) it is a priori that whatever has Φ has Ψ .

Premise three states that human beings have qualia. This is a generic, like "an eagle has wings." To say that human beings have qualia is to say that normal human beings have certain types of subjective experiences. Of course, in completely sound, dreamless sleep, a human being has no subjective experiences, and so no qualia. But premise three, in the intended sense, implies only that a normal human being is able to have qualia, that is, is able to have subjective experiences.

Turn to premise four. In the second sentence in the premise, "All properties of a complex system that are not relations between it and something else derive from the properties of its constitutes and their effects on each other when so combined," I'll take "a relation between it and something else" to mean a relation between it and something wholly distinct from it, and so something distinct from it that is also not a part of it. The second sentence is Nagel's gloss on what he means by denying that there are emergent properties. He takes the two claims in premise four to be equivalent.

For the argument to be valid, the notion of property derivation in premise four must be the notion of logical property implication in premise two. A property Ψ is derivable from a property Φ just in case Φ logically implies Ψ . Thus, to say that a property Ψ of a complex system derives from the properties of its constituents and their effects on each other when so combined is to say that it is logically necessary that any system with constituents propertied in the ways in question, and effecting each other in the ways in question, has Ψ .

There is a trivial way that any property of a complex system will be logically implied by properties of its constituents and their effects on each other when so combined (Van Cleve 1990, 223). If c is a constituent of a system S that has Ψ , then c has the property of being a constituent of a system S that has Ψ . That property of c logically implies Ψ . Also, one of the effects that the constituents have on each other when so combined is forming a system that has Ψ . That relational property logically implies Ψ . If the conclusion is to follow from the premises, premise four must be revised.

It will help here to appeal to a notion of a microstructural property. Let's say that a microstructural property is a property of being wholly decomposable into nonoverlapping entities that are respectively *intrinsically* propertied in soand-so ways, and have such-and-such effects on each other when so combined, where the effects include only changes in intrinsic properties of constituents and/or relations that the constituents bear only to each other.⁶ Microstructural properties can include intrinsic dynamic properties. A neuron's firing, for instance, is an intrinsic dynamic property of the neuron. Moreover, the relations among constituents involved in the microstructural property can involve changes in the intrinsic properties of constituents. A microstructural property of a system isn't a property that the system has just in virtue of the constituents forming a mereological sum, or just in virtue of their having a certain spatiotemporal arrangement. Certain microstructural properties will be such that something counts as a system in virtue of having them.

Armed with this notion of a microstructural property, we might replace premise four with the following assumption:

Premise Four*. There are no emergent properties. That is to say, every intrinsic property of a complex system is logically implied by some microstructural property of the system.

The second claim is supposed to be equivalent to the first. In all the leading conceptions of property emergence of which I'm aware, the second claim implies the first (for intrinsic properties). But given that the first is supposed to imply the second, Nagel has a truly unusually liberal notion of property emergence in mind.⁷ Of course, 'emergence' is a philosophical term of art. Nagel is within his rights to stipulate that in his sense of an emergent property, an intrinsic property of a system is an emergent property if it fails to be logically

implied by any microstructural property of the system. The problem is not with his use of 'emergence.' He can drop the first sentence of premise four, and we can drop the first sentence of premise four*. What is intended to do the work in the argument is the claim that every intrinsic property of a system is logically implied by some microstructural property of the system.

But premise four^{*} is inadequate. If a complex system is decomposable at more than one scale, it will have microstructural properties at different scales. Premise four^{*} leaves open whether there are scales at which S is wholly decomposable into constituents that don't figure in any microstructural property that logically implies any of the system's contingent intrinsic properties. Putting the matter again in our terminology, Nagel intends premise four to imply that for any intrinsic property Ψ of a system, and any scale at which the system is wholly decomposable into nonoverlapping entities $c_1 \dots c_n$, there is some microstructural property involving entities $c_1 \dots c_n$ that logically implies Ψ . The microstructural property will be such that anything that has it thereby constitutes a system of the type S in question. Premise four is intended to have that consequence. Premise four^{*} doesn't.

For the purpose of making a case of panexperientialism, it needn't be assumed that every proper kind of constituent⁸ of a type of system S will figure in some microstructural property in virtue of which a system of type S is constituted and that logically implies every intrinsic property of any S so constituted. It need be assumed only that every quale of any system so constituted is so logically implied. I'll thus replace premise four with the following assumption:

Premise 4. Every kind of proper constituent of a type of system S with qualia will figure in some microstructural property in virtue of which a system of type S is constituted and that logically implies all of the qualia of any S so constituted.

Premise 4 implies that qualia are intrinsic properties of systems that have them. The reason is that no contingent extrinsic property of a system will be logically implied by any of its microstructural properties. But panexperientialists maintain that qualia are intrinsic. Of course, some of the currently leading contenders for the correct theory of phenomenal consciousness entail that qualia are extrinsic properties (see Dretske 1995; Tye 2000; Hill 2009). The premise thus requires defense. Although I, myself, am not a panexperientialist, I think that qualia are intrinsic properties in the sense in question. I think that any intrinsic duplicate of a phenomenally conscious being would itself be phenomenally conscious. So, I, myself, won't challenge premise 4 on the grounds that the qualia of a complex system are extrinsic properties of it. Let it suffice for me to note, moreover, that I don't think that taking qualia to be relational properties will change of any the basic issues that we'll be discussing in what follows. Instead of invoking microstructural properties of systems, we could instead invoke microproperties of systems that include relations the system bears to things that are wholly distinct from it; indeed we could invoke microproperties of human beings that involve relations to everything else in the world. The same basic issues I'll be discussing would still arise.

Even if premise four is replaced with premise 4, the conclusion still doesn't follow. Premise one ensures that the microstructural properties of a human being will include only physical constituents. But one way that a microstructural property could still fail to be physical, or at least wholly physical, despite including only intrinsic physical properties, is if one or more of the relevant relations among the constituents fails to be physical.

I propose that we handle this in our characterization of a microstructural property. Nagel appeals to the relation of being-a-constituent-of-something and the relation of one property logically implying another. They are both topic-neutral relations. The causal relation is also a topic-neutral relation. It won't matter for present purposes whether we count spatiotemporal relations as physical or instead as topic neutral (on the grounds, say, that being in spacetime doesn't logically imply being physical), so long as we count them as one or the other. Let's say that a microstructural property is a property of being wholly decomposable into nonoverlapping entities respectively intrinsically propertied in so-and-so ways, and having such-and-such causal effects on each other when so combined, where the effects include only changes in intrinsic properties of constituents and/or relations that the constituents bear only to each other, and where (here's the new qualification) the relations are only physical and/or topic-neutral relations. I'll count a microstructural property as physical if the intrinsic properties it involves are all physical properties and the relations it involves are all physical or topic-neutral ones. Given premise two, then, if a microstructural property of a system logically implies a quale, then at least some of the intrinsic properties it involves are nonphysical properties.

Gaps in the argument remain. One is from the claim that the basic physical constituents of the universe have nonphysical properties to the claim that basic physical constituents of the universe have qualia. Another is from the claim that some basic physical constituents of the universe have qualia to the claim that all basic physical constituents of the universe have qualia. I'll discuss these gaps in reverse order.

By 'the basic physical constituents of the universe' in the conclusion is meant *all* the basic physical constituents of the universe. The 'all' is necessary for the 'pan' of panexperientialism. The premises at best imply only that at least some of the basic physical constituents of the universe have qualia.

Van Cleve (1990, 216) thinks that Nagel succeeds in closing this gap in the argument when he asserts: "Anything whatever, if broken down far enough and rearranged, could be incorporated into a living organism. No constituents beside matter are needed" (Nagel 1979, 181). Premise one gives us that humans are wholly constituted by physical particles, and so no constituents beside matter are needed. It could turn out, moreover, that anything whatever, if broken down far enough and rearranged, could be incorporated into a living organism. But that additional contentious empirical assumption won't get us to the conclusion that all the basic constituents of the universe have nonphysical properties. The reason is that it doesn't follow from that assumption that every basic constituent of matter is such that organizations of it alone could constitute a human being. That would follow if there were a single kind of basic constituent of matter. But, on the evidence, there isn't. Suppose that electrons are indeed basic physical constituents of the universe. No arrangement of electrons alone could constitute a human being, or any living organism, or indeed even any chemical element. The premises won't establish that electrons have qualia, and so won't establish that all of the basic constituents of the universe have qualia.

I propose that we close this gap by weakening the conclusion to the claim that *at least some* of the basic physical constituents of the universe have qualia. Although that doesn't deserve of the name 'panexperientialism' (or 'pansubjectivism' or 'panpsychism') since it allows some basic constituents of the universe to fall through holes (though hopefully not the relevant wholes) in the pan, it is a very closely related doctrine that is of course of enormous interest if true. Even if not all particles of matter are mind dust, it is, needless to say, of tremendous interest if some are. If true, there would be, in addition to the completely misleadingly labeled 'God particle,' accurately labeled 'conscious particles.' There would be mind dust. Phenomenal consciousness, and so subjectivity, would be present at the most fundamental physical level.

Let's turn to the second gap I mentioned. The premises don't jointly imply that at least some basic constituents of the universe have qualia. At best, they imply that at least some basic constituents of the universe have intrinsic nonphysical properties. It is consistent with the premises that mental properties are logically implied by microstructural properties that include intrinsic properties of constituents that are neither mental properties nor physical. As Van Cleve (1990, 217) points out, Broad (1925) called that view 'neutralism.' Perhaps there are relevant intrinsic properties of certain physical objects that are neither physical (even in the broadest sense) nor mental.

As Van Cleve (1990, 217) also points out, Nagel tries to close this gap by stipulation, saying: "This [view] could still be called pansychism" (Nagel 1979, 185). Van Cleve spots Nagel that claim. The reason, I suspect, is that

even though Nagel states the conclusion of his argument as the claim that the basic constituents of the universe have mental properties, he actually intends to establish the conclusion that the basic constituents of the universe have what he, Nagel, calls 'protomental properties' (Nagel 1974/1979; 1989; 1999). As Nagel's various discussions of protomental properties make clear, protomental properties are supposed to be neither mental properties nor physical properties. Indeed, they are supposed to figure, in conjunction with physical properties, in reductive explanations of mental properties.

The doctrine that the basic physical constituents of the universe have protomental properties shouldn't be thought of as a kind of pansychism. The reason is twofold. First, protomental properties aren't mental properties. Second, on any doctrine deserving of the name 'pansychism,' mental properties (or at least some mental properties) should be fundamental properties, properties that cannot be reductively explainable. The view that at least some basic constituents of the universe have protomental properties is a kind of neutralism. But the kind in question should be distinguished from neutral monism, the view that both mental properties and physical properties are logically derivable from properties of a single, more fundamental, kind that is neither mental nor physical. (I won't pause to discuss neutral monism here, though some of what I say below will be directly relevant to that doctrine.) The idea is, rather, that mental properties are possessed just in virtue of possessing certain physical and protomental properties. This view is best called 'panprotopsychism,'9 rather than 'pansychism.' Nagel's panprotopsychism stands in opposition to pansychism and panexperientialism in denying that any mental properties are fundamental properties. But unlike neutral monism, panprotopsychism allows that certain physical properties are also fundamental properties, and thus doesn't entail that the physical is reducible to something else.

We are concerned with panexperientialsim, rather than the more general doctrine of panpsychism. The relevant protomental properties where qualia are concerned are protoqualia. Protoqualia would be different from qualia in that they could be possessed by something that is not a subject of experience, that has no subjective perspective on the world, that is devoid of subjectivity. Protoqualia are supposed to be such that they can figure along with physical properties in reductive explanations of qualia and also of the property of being a subject of experience. Panprotoexperientialism thus doesn't imply pansubjectivism. Indeed, it entails that no basic physical constituent of the universe is a subject of experience. I'll return to panprotoexperientialism in due course. For now let's continue to focus on panexperientialism.

Even with the revisions we've made, the premises of the argument still at best imply only that at least some basic physical constituents of the universe have nonphysical properties. To get an argument for the conclusion that at least some basic constituents of the universe have qualia, we need to strengthen premise two.

We could so as follows: Qualia are not logically implied by any other properties. That rules out that qualia are logically implied by certain combinations of physical properties and protoqualia. That premise, however, is too strong. At least it is too strong if we embrace an abundant theory of properties. On an abundant theory, there are disjunctive properties and negative properties. Consider, then, any quale Q and any contingent physical property P. There is the disjunctive property (Q or P) and the negative property not-P. Neither property is it itself a quale. The disjunctive property (Q or P) could be possessed by something entirely devoid of qualia, in virtue of its having P. Not-P too could be possessed by something entirely devoid of qualia. What can't be possessed by something entirely devoid of qualia is the conjunctive property (Q or P) and not-P. That property logically implies Q: it is logically necessary that whatever has (Q or P) and not-P, has Q. Thus, a quale will be logically implied by properties that are not qualia.¹⁰

Qualia aren't closed under Boolean operations. Thus, for instance, not-Q isn't a quale. Although the disjunctive property (Q or P) is not a quale, it includes a quale as a constituent, namely Q. Let's call a property 'wholly nonqualitative' just in case it is not a quale and includes no quale as a constituent. (Q or P) fails to be wholly nonqualitative. We can then replace premise two with the premise:

Premise 2. No quale is logically implied by any wholly nonqualitative property.

I briefly mentioned earlier an issue about premise one. It may be that the basic constituents of the universe are not particle-like but rather strings of vibrating energy, or branes, or something of some other sort. There is also an empirical question of whether there is a bottom level of the physical universe, a question hasn't been conclusively settled. Pansychism and panexperientialism shouldn't take a stand on such issues. I suggest that we handle this by replacing premise one with the following assumption:

Premise 1. Human beings are complex systems wholly composed of atoms.

Given this replacement, the conclusion that at least some basic constituents of the universe have qualia doesn't follow. But I suggest that we draw the following conclusion instead from the premises:

Conclusion. At least some of the subatomic physical entities into which atoms are wholly decomposable have qualia; and at any level at which
those subatomic entities are themselves wholly decomposable, they have some physical entities as constituents that have qualia.

This conclusion leaves open whether every physical entity at or below the scale of human beings has qualia. And it leaves open whether physical entities at the same scale or at larger scales have qualia. Still, we could count the above conclusion as expressing a kind of panexperientialism. If anyone objects on the ground that the doctrine doesn't entail that everything (even everything that is spatiotemopral) has qualia, we could instead call the doctrine expressed by the conclusion 'subatomic-experientialism.' If it is true, there's mind dust.

It could well turn out that there is a smallest physical scale. A Planck length may be such. Moreover, should there turn out to be basic physical constituents of the universe (whether they are particle-like or not), and so a bottom physical level, then that fact, in conjunction with the above conclusion would imply that at least some basic physical constituents of the universe have qualia; and, so, imply, as well, that at least some basic physical constituents of the universe are subjects of experience. Given subatomic-experientialism, then, the issue of whether at least some basic physical constituents of the universe are subjects of experience just turns on whether the universe has basic physical constituents.

I'll now present the argument we've been building for panexperientialism or subatomic-experientialism (repeated premises are renamed):

- Premise 1. Human beings are complex systems wholly composed of atoms.Premise 2. Qualia are not logically implied by any wholly nonqualitative properties.
- Premise 3. Human beings have qualia.
- **Premise 4**. Every kind of proper constituent of a type of system S with qualia will figure in some microstructural property in virtue of which a system of type S is constituted and that logically implies all of S's qualia. **Lemma**. At least some atoms have qualia.
- **Conclusion**. At least some of the subatomic physical entities into which atoms are wholly decomposable have qualia; and at any level at which those subatomic entities are themselves wholly decomposable, they have some physical entities as constituents that have qualia.

The conclusion is, to put it mildly, hard to swallow. One might declare it absurd and take the argument to yield a *reductio* of the premises. But people living in the glass house of a theory of phenomenal consciousness should be careful about throwing stones like 'absurd.' We can leave those stones to people that sit on the sidelines, who opt out of trying to give a theory of phenomenal consciousness. Those of us in the game should be civil with each other. So far as I can determine, the conclusion isn't a priori false. It seems coherently conceivable, for instance, that an elementary physical particle is a subject of consciousness, that it has some kind of 'internal phenomenal buzz.' I, myself, at any rate, know of no a priori necessary condition for being a subject of experience that would rule that out. The above argument doesn't present a paradox. Still, one might think there isn't even the slightest reason to believe that some atoms and subatomic entities have qualia. But there is reason to believe it. The above argument offers reason to believe it. To be sure, the reason is defeasible since, by everyone's lights, the argument includes a posteriori premises. Still, the premises enjoy some credibility, and they jointly lead to the conclusion.

I suspect that many will feel that whatever credibility the premises enjoy, the conjunction of the premises enjoys less credibility than the denial of the conclusion. In what remains, I'll leave the odd premises out, and focus just on whether we should accept the conclusion or instead deny either premise 2 or premise 4.

Panprotoexperientialism entails the denial of premise 2. It entails that there are wholly nonqualitative properties that logically imply qualia. The properties in question will include at least some protoqualia.¹¹

Must physicalism entail the denial of premise 2? I'll take it that any doctrine deserving of the name 'reductive physicalism' will deny it. Reductive physicalism implies that qualia can be reductively explained just in physical (and topic-neutral) terms alone, and so entails that premise 2 is false. It does so, because an epistemic condition on reductive explanation is that the *explanans* logically implies the *explanandum*.

Emergentism as concerns qualia entails premise 2. That is, it entails that there are no wholly nonqualitative properties of any sort that logically imply qualia. On that score, emergentists and panexperientialists are in agreement. But emergentism as concerns qualia entails the denial of premise 4. (I'll say more about property emergence shortly.)

Assuming premises 1 and 3, if we reject panexperientialism, we *seem* to be left with a stark choice. Either there are wholly nonqualitative properties that logically imply qualia or qualia emerge from wholly nonqualitative properties. But I'll argue later that there's another option.

Panprotoexperientialists and reductive physicalists incur the dialectical obligation of making a compelling case that there are wholly nonqualitative properties that logically imply qualia. As should be apparent, that dialectical obligation is truly formidable. Panexperientialists maintain it can't be discharged.

Emergentists incur the dialectical obligation of making a compelling case that qualia are properties of certain complex systems, despite not being logically implied by any wholly nonqualitative properties of the system. But panexperientialists incur that dialectical obligation too. It is common ground between emergentists and panexperientialists. The emergentist rejects only premise 4. Here the burden is on the panexperientialists to show that all the qualia of a complex system are logically implied by some microstructural property of the system. Emergentists maintain it can't be discharged.

Panexperientialists, panproperexperientialists and physicalists take emergentists to be mystery mongering. Let's consider that charge. In a famous passage, Huxley remarked: "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 Djin when Aladdin rubbed his lamp" (Huxley 1974, 192). It is also remarkable that when a match head is rubbed across a rough surface, a flame appears. But the flame's appearing when the match is rubbed is, we now know, accountable: there is an underlying mechanism that explains how the match's being rubbed results in the flame. By hypothesis, there is no underlying mechanism that explains how the lamp's being rubbed results in Djin appearing. It just happens, and there is an end to it. According to the emergentist, there is no underlying mechanism that explains how a microstructural property results in a state of phenomenal consciousness. It just happens, and there is an end to it. Djin's appearing when the lamp is rubbed is magic. It seems that the emergentist too is committed to magic. Qualia magically arise from microstructural properties. Somehow, magically, from the one microstructural property alone we get another distinct property not logically implied by it.

Emergentists don't, however, just deny that qualia are logically implied by any wholly nonqualitative properties. They make a positive claim too. I'll focus here on a notion of intrinsic property emergence due to Broad (1925).¹² The notion can be explicated as follows (see McLaughlin 1997):¹³

An intrinsic property Ψ of a complex system S is an emergent property of S if and only if no microstructural property of S logically implies Ψ , but some microstructural property Φ of S is such that it is a fundamental law of nature that whatever has Φ has Ψ .

A law is a fundamental law in the intended sense just in case it does not hold just in virtue of other laws and conditions. A fundamental law is both ontologically and epistemically fundamental. It is thus an irreducible law. The property of being a fundamental law is a global property: it depends on the way the rest of the world is. It depends on what other laws and conditions hold in our world.¹⁴ The first conjunct of the analysans implies that the fundamental law of nature in question is a contingent law. On this view, a property Ψ of a complex

system S emerges from a property Φ of S if Φ is a microstructural property of S and it is a contingent, fundamental law of nature that whatever has Φ has Ψ . Φ 's emerging from Ψ thus also depends on a global property, namely on the law linking them being a fundamental law of our world.

Notice that given that the laws linking physical microstructural properties with qualia are fundamental laws, there is "a zombie world." That is to say, there is a possible world that is an exact physical duplicate of our world yet is entirely devoid of qualia, and so a world in which our physical duplicates are 'zombies' (see Chalmers 1996).

To return to Aladdin's lamp, the emergentist will claim that Djin's appearing when the lamp is rubbed is indeed magic if there is no general law of nature at work. Suppose, however, that we discover, through empirical investigation, a law of nature according to which if a lamp of a certain kind is rubbed in a certain manner, then a djin (a djin-like being) will appear. We could then explain why a djin appeared when a certain lamp was rubbed by citing the law, and pointing out that the lamp is of the kind in question and was rubbed the way in question. That would render the djin's appearance within the realm of the natural, rather than the supernatural; it would be covered by a law of nature. It would thus be consistent with naturalism. Still, though, it would no doubt continue to seem mysterious why it is that a djin appears whenever a lamp of the kind in question is so rubbed. Why, we'll wonder, when such a lamp is rubbed in such a manner does a djin appear, rather than something else, or nothing at all? We'll feel that there should be some underlying mechanism that explains how the rubbing of such a lamp gives rise to a djin appearing. (The appearance of a flame when a match head is rubbed would have continued to seem mysterious too, had we found no underlying mechanism. Why does a flame appear when the match head is rubbed, rather than something else, or nothing at all? That, it turns out, is a question we can answer.) If, however, the law of nature is fundamental, then there is no such mechanism to be found. The emergentist will say that we must, then, simply accept, with what Broad and Alexander called 'natural piety' that that's just what happens when such a lamp is rubbed in such a way. If that should still be regarded as magic, then the law in question is a law of magic. The fundamental laws of our world include laws of magic; magic is part of the natural world. If fundamental laws linking physical mircostructural properties and qualia would be laws of magic, then there's magic in our world. We should be enchanted by it.

Nagel (1979, 286–87) remarks that we should expect to find correlations between qualia and brain states. Suppose we find that whenever a human being's brain is in mircostructural state B, the human being is in a state of phenomenal consciousness C. Suppose that we find that that generalization is strictly true and counterfactual supporting. We'll then face what Chalmers calls 'the hard problem of consciousness' (Chalmers 1995; see also Chalmers 2002, 248): Why does B give rise to C (e.g., the feeling of pain) rather than to some other state of consciousness (e.g., the feeling of gentle warmth) or no state of phenomenal consciousness at all?¹⁵ If the psychophysical generalization expresses a fundamental law of nature, and thus admits of no explanation, then there is no answer. There will be no mechanism that explains how C arises from P. Indeed, C arising from P will be inexplicable, a fundamental law of nature. C *emerges* from P. We can do no more than accept that with natural piety.

Many find it wildly implausible that there are such fundamental psychophysical laws. Herbert Feigl claimed that emergentism entails that qualia are 'nomological danglers.'¹⁶ The problem, though, isn't just the grotesquely baroque view it entails of the nomological structure of reality, a view according to which we couldn't write the fundamental laws of nature on a T-shirt in big enough type to read without an electron microscope. That's not the main problem. I'll simply mention the main problem since it's now well-known (and a problem that worried Broad). The main problem is that physical theory seems to have no need of the hypothesis that there are qualia.¹⁷ Given that, and given that qualia are not ontologically derivative from any of the phenomena that physical theory posits to explain motion, it seems that there is no way for qualia to make a difference to causal relations. If so, that has the truly unhappy consequence that the feel of pain, for instance, makes no causal difference to anything, not even one's belief that one is in pain.¹⁸ Emergentism thus seems committed to an abhorrent kind of property or type epiphenomenalism, one that leads to token epiphenomenalism on a property exemplification account of states and events.¹⁹

Panprotopsychists and physicalists find these consequences unacceptable on grounds of overall simplicity and coherence. They deny that qualia emerge from physical properties. Panprotopsychists and reductive physicalists maintain that there are wholly nonqualitative properties that logically imply qualia; that for every quale, there is some wholly nonqualitatve property that logically implies it. Panexperientialists maintain that there are microstructural properties of complex systems that logically imply qualia in that sense. But they deny that any of those microstructural properties are wholly nonqualitative. They maintain that they will include qualia as constituents. Let's consider first panexperientialism, then panprotoexperientialism, and then physicalism.

Panexperientialism faces a host of problems that have long been recognized. First, it doesn't solve the general mystery of how something physical could have qualia. As Nagel (1999) noted, on the panpscyhist view, there is qualia/physical property dualism all the way down—all the way down the micro-macro scale. That was recognized both by William James (1890/1950) and by Arthur Lovejoy (1927). Suppose there are fundamental physical particles, electrons, say, that have a certain state of phenomenal consciousness C. We then face the hard problem for electrons: Why do electrons have C rather than some other state of phenomenal consciousness C^* or no phenomenal consciousness at all? The panexperientialist will have to say that we have to accept with natural piety that electrons have C.

Of course, panexperientialists think that qualia are fundamental and so cannot be reductively explained. They maintain, however, that we can at least explain how complex systems have qualia by appeal to certain microstructural properties of the systems that include qualia of the system's constituents. But that position faces a problem that has long been recognized. As Van Cleve aptly notes, the view makes us out to be made up of conscious homunculi (Van Cleve 1990, 220). We have no compelling examples of how a group of entities with qualia can become so related that they constitute a system that itself has qualia. We know we have qualia and can participate in systems of various sorts. But we have no compelling examples of how our having qualia results in a system we partly form having qualia. We say things like, "The team feels exhausted." But we take that to be a matter of the relevant members of the team feeling exhausted. We don't take there to be an entity, the team, over and above the members, that itself has the feeling of exhaustion. Also as concerns our own qualia, it seems it doesn't require a village of homunculi for us to feel pain. Indeed, even if there were a village of homunculi inside us feeling pain, we don't see how that could amount to our feeling pain. The problem that the panexperientialist faces, as Van Cleve notes, is how one would "account for the unity of consciousness that somehow arises out of the colony of homunculi]" (Van Cleve 1990, 220; emphasis his). This problem too was recognized by William James (1890/1950), and now goes by the name 'the composition problem' (Coleman 2012). Let it suffice to note that it is a truly formidable problem indeed.

Moreover, panexperientialism raises a slew of new mysteries besides how conscious beings can make up a conscious being. Consider atoms. Which of them have qualia? All of them or only some of them? If only some of them, which? What sorts of qualia do they have? Do they have the sorts of qualia we have? Is there, for instance, an atom that has the feel of pain? If so, does it always have it, or does it sometimes have it and sometimes not? Is there another atom that has the feeling of gentle warmth? Or does a single atom have both the feel of pain and the feel of gentle warmth? Could a single atom have both the feel of warmth and the feel of bitter cold in different parts? Are there atoms that have visual experiences or that have auditory experiences? If so, are they accurate enough for them to in some sense see and hear?

Perhaps atoms and more fundamental physical entities have qualia but just not any of our qualia. Perhaps they have entirely alien qualia.²⁰ If the qualia are entirely alien, then they have no determinable quale that we also have. The idea is not incoherent. There seems to be no super-determinable quale. It's like something to have a stabbing pain in one's toe, like something to see a rainbow, like something to smell a fragrant odor, and like something to feel dizzy. But I, for one, can discern no determinable quale of which the qualia in question are all determinables. As Nagel (1974/1979) notes, a bat has qualia when it perceives by echolocation that we don't have; and the sighted among us have qualia that the blind among us don't have. It seems that something could have qualia that are entirely different from our qualia, completely alien qualia. But if atoms and more fundamental physical entities have qualia that are entirely alien, then we can have no empathetic understanding whatsoever of what it is like to have their qualia. What it is like to be, say, a hydrogen atom would be in no way whatsoever like what it is like to us. We could have no conception of what it is like to be a hydrogen atom. We would be cognitively closed to such qualia. How can alien qualia, qualia such that we can have no conception whatsoever of what it is like to have them, figure in explanations of our qualia? Well, perhaps they can. But we would be cognitively closed to the explanations. They'd be beyond our ken.

Of course, every extant theory of phenomenal consciousness faces truly formidable problems. But the problems faced by panexperientialism and emergentism seem to me sufficiently severe that we should look elsewhere for a theory of phenomenal consciousness.

Both panprotexperientialism and reductive physicalism entail that there are wholly nonqualititave properties that logically imply qualia. Panprotoexperientialists maintain that we can see a priori that no physical property logically implies any quale, and so that reductive physicalism is false. They use arguments such as the knowledge argument and conceivability arguments to show that. They also maintain that panexperientialism and emergence are wildly implausible for just the sorts of reasons given above. They claim that there must be properties of a kind we've yet to discover that, in conjunction with physical properties, logically imply qualia.

Van Cleve notes at one point in connection with Nagel's appeal to protomental properties:

One could . . . [claim] that there may be logical connections between proto-mentality and full-blown mentality, even if we are at present incapable of discerning them. In general, one can always raise the possibility than an apparent case of emergence is merely apparent, owing to the existence of logical connections not yet fathomed by us. But this would be to raise at the same time the possibility of unfathomed logical connections between the physical and the mental, thus undermining the case for premise 2 [our premise two]. Defending the argument in this way would only lay it open to attack somewhere else. (Van Cleve 1990, 220)

One issue is whether there are not just now unfathomed logical connections between the physical and qualia, but logical connections that are unfathomable. One could try to reject premise two and premise 2 on the grounds that there are unfathomable logical connections between the physical and the mental that amount to a reduction of the mental to the physical. That is Colin McGinn's (1989; 2001) position; a position that's been labeled 'mysterianism.'²¹ McGinn (2001) maintains that mental properties reduce to physical properties, but that we are cognitive closed to the physical properties such a reduction would have to invoke. Notice that requires that the physical properties in question are very different indeed from the physical properties that we are familiar with even in the most exotic reaches of physical theory. They are so different that we've cognitively closed to them; they run against the grain of our thought. We can't conceptualize them.

If atoms and more fundamental physical entities have entirely alien qualia, and so qualia that are for us inscrutable, then panexperientialists too are committed to the view that our qualia reduce to properties to which we are cognitively closed, and so we are cognitively closed to the explanations. McGinn's brand of physicalism and this brand of panexperientialism (according to which atoms and more fundamental physical entities have alien qualia) both entail that the qualia of complex systems are reductively explainable, but we are cognitively closed to the explanations, because we're cognitively closed to the *explanans*. The explanations will have to cite properties we can't get our minds around.

Panprotoexperientialists deny that qualia reduce to physical properties, fathomable or not. But they maintain that they are reducible, not fundamental. Unlike mysterians of any ilk, they're not pessimistic about our conceptual prospects for discovering the properties needed for the reduction. They hold that while we have yet to discover any protoqualia, there is good reason to think there are such properties, and no good reason to think we're cognitively closed to them. They maintain that we could with time, effort, luck, and enough money someday discover protoqualia. Protoqualia, they maintain, are so linked to qualia that certain microstructural properties involving protoqualia and physical properties (and topic-neutral relations) logically imply qualia. As I noted, this is Nagel's view in "Panpsychism" (Nagel 1979). It is a view that appears to be gaining adherents.

There are senses of the term 'physical' in which protoqualia, as the panexperientialist conceives them, would be physical properties. For instance, if a property is posited by the physics that is in fact true of our world, then it is a physical property. At least some panexperientialists are betting that the physics true of our world will have to posit protoqualia. In the sense of 'physical' in question, they'd then be physical properties. But that's not the relevant sense of 'physical' in the debate between physicalists and panexperientialists. There is an interesting question of what is meant by 'physical' in that debate; there is no consensus about that (cf. Montero 2009). One of the leading ideas, though, is that a property is a physical property if it is posited by current physical theory or by some recognizable descendent of current physical theory (see Jackson 2007). A panexperientialist might maintain that protoqualia are a kind of property that is neither a vernacular physical property, nor posited by current physical theory, nor a kind that would posited by any recognizable descendent of current physical theory. The idea, which is admittedly quite vague, is that the physics that posits protoqualia will be sufficiently different from current physics that it will not be a recognizable descendent of it. How different it might be from current physics is an interesting open question. But one idea is that the future physics that posits protoqualia could be as different from current physics as current chemistry is from air, earth, fire, and water theory (an analogy due to Nagel).

Given the flexibility of the term 'physical,' panprotoexperientialists needn't insist that there is no sense of 'physical' in which protoqualia will count as physical. The important difference between the kinds of physical properties with which we are familiar from everyday discourse and from physical theory, on the one hand, and protoqualia on the other hand, is supposed to be this: protoqualia can combine with physical properties (in the familiar sense) to logically imply qualia; indeed, to reductively explain qualia. That characterization would leave open that they are the physical properties McGinn had in mind (to the extent that he had anything definite in mind). The dispute between mysterians such as McGinn and panprotoexperientialists would then be over whether there is sufficient reason to think that we are cognitively closed to protoqualia.

The issue of what counts as physical aside, protoqualia are supposed to have the following features: (1) they are intrinsic properties of at least some subatomic physical entities; (2) they are ontologically fundamental; (3) they are neither qualia (or any other kind of mental property) nor physical properties (in the familiar sense); (4) they figure along with intrinsic physical properties involved in microstructural properties that constitute complex systems with qualia; and (5) the microstructural properties in question logically imply the qualia of the complex system in question. Although we know (1)-(5) are supposed to be true of protoqualia, we cannot now see, even as through a glass dimly, what these properties are. Mysterians, such as McGinn, think we are cognitively closed to them. Panprotoexperientialists deny that.

Nagel famously remarked: "Consciousness is what makes the mind-body problem really intractable" (Nagel 1974/1979, 165). McGinn (1989) thinks it is so intractable that we could never solve it, though we can somehow see that it must have a physicalist solution, just one that will forever fall outside our ken. Panprotoexperientialists think the problem is really intractable, but that we can in principle solve it since we can in principle discover protoqualia, properties that meet conditions (1)-(5). (Of course, though, a kind of mysterian panprotoexperientialism is a position in logical space.) Physicalists think we can solve it without any truly revolutionary change in our theory of the physical world.

Nagel, I believe, pinpointed why phenomenal consciousness makes the mind-body problem really intractable. I'll state the matter bluntly, avoiding necessary but obscuring qualifications. In a nutshell, the reason it does so is that qualia are subjective and physical (and functional) properties are objective. He tells us: "A feature of experience is subjective if it can in principle be fully understood only from one type of point of view: that of a being like the one having the experience, or at least like it in the relevant modality" (Nagel 1979, 188). We'll put the point this way: one can fully understand a quale only if one knows what it is like for a subject to have it. One can't fully understand the feel of pain, for instance, unless one knows what it is like for a subject to feel pain. Physical (and functional properties), by contrast, are objective properties.

What is deeply puzzling about the notion of protoqualia is that they are supposed to build an a priori bridge between subjective properties such as qualia and properties that are not subjective. It is claimed that physical properties cannot logically imply qualia, because physical properties are objective properties and qualia are subjective properties. But protoqualia are not qualia. It's not like anything to have them. That's why having them doesn't require being a subject of experience. It seems that they're objective. But it is, then, deeply obscure what their status is supposed to be such that combinations of them and physical properties, wholly nonqualitative properties, logically imply qualia.

We can run a knowledge argument in terms of objective truths (see McLaughlin 2003a). Suppose that LaPlace's demon can't feel pain or have any qualia like the feel of pain. LaPlace's demon can know all the objective truths of our world and that they are all of the objective truths of our world. But that knowledge will not yield knowledge of what it is like to feel pain.

Also, for any objective property, it seems we'll be able to imagine having it without having any quale. The reason is that we'll be able to imagine having the objective property without having to empathetically imagine having a quale (see Hill and McLaughlin 1999).

Nagel (1974/1979; 1979; 1989) sometimes writes as if protoqualia (protomental properties) will have some kind of intermediate status between being subjective and being objective. But given his own characterization of that distinction, there is no intermediate status. We can characterize degrees of objectivity, and corresponding degrees of subjectivity; but every property, it seems, will fall somewhere within that spectrum. Protoqualia are supposed to be such that a being can have them even though there is literally nothing it is like to be that being. That makes them objective. But there seem to be no a priori links between subjective and objective properties. Perhaps there are protoqualia, but we are cognitively closed to them. Or perhaps, as Nagel (1989) has suggested, we are in the position of a pre-Socratic being told that matter is energy. We just now happen to lack the concepts to make any sense of that claim.

I admire the optimism of panprotoexperientialists. But the point I want to underscore is that panprotoexperientialism seems to me to be motivated just by a prejudice. The prejudice is that qualia cannot be epistemically fundamental (perhaps because they are possessed only by certain very complex systems), that they must somehow be reductively explainable. They are, the panprotoexperientialist claims, not reductively explainable by physical (and functional properties), so there must be some other sort of property that when combined with physical (and functional properties) reductively explains them. Perhaps, however, qualia are, as they seem to be, epistemically fundamental, despite being possessed only by certain complex beings, including ourselves. As we'll see, if they're epistemically fundamental, it wouldn't follow that they are ontologically fundamental.

I, myself, accept premise 2 of the earlier argument for panexperientialism. Qualia are not logically implied by any wholly nonqualitative properties. I thus reject reductive physicalism (see McLaughlin 2007). There is, I think, indeed an unbridgeable explanatory gap between qualia and the physical-functional (see Levine 1983; 2001); indeed, between qualia and the wholly nonqualitative. I accept premises 1, 2, and 3. I deny only premise 4. So, as concerns the argument, I have the same response as an emergentist. I'm not, however, an emergentist as concerns qualia in anything like Broad's sense. The reason is that I don't think there are fundamental laws of nature linking any physical mircostructural properties of complex beings with qualia, and the qualia of those beings. I'm a kind of epistemic emergentist, rather than an ontological emergentist. In what remains, I'll lightly sketch the view I favor. But first some quick preliminaries.

Logical implication, you'll recall, is both a metaphysical and epistemic notion. Let's separate the two aspects. Let's say that Φ metaphysically implies

 Ψ just in case it is metaphysically necessary that (true in literally every possible world that) whatever has Φ has Ψ ; and let's say that Φ epistemically implies Ψ if and only if it is a priori that whatever has Φ has Ψ (see Chalmers and Jackson 2001). As Saul Kripke taught us, neither requires the other. He famously showed that a statement can fail to be a priori yet nevertheless be metaphysically necessary (see Kripke 1971; 1980). The story is well-known, but worth briefly repeating. Kripke (1971) gave an a priori argument for the necessity of identity thesis, the thesis that if A = B, then necessarily A = B. He derived that thesis a priori from two other a priori theses: the thesis that everything is such that it is necessarily identical with itself, and the thesis that if A = B, then whatever is true of A is true of B. He allowed that there are contingent statements of identity, such as "Benjamin Franklin is the inventor of bifocals," even though identity itself is necessary. We get a contingent statement of identity when one of the terms flanking the identity sign fails to be a rigid designator, that is, fails to designate the same thing in every possible world in which it designates anything. When, however, the terms flanking the identity sign are rigid designators, then the identity statement will be noncontingent (in the modally strongest sense). As Kripke notes, even though it was an empirical discovery that water is (is identical with) H₂O, given the necessity of identity (and the fact the kind term 'water' and the descriptions 'H₂O' are rigid designators), the statement "water is H₂O" is necessarily true; necessarily true, despite being a posteriori. The statement that there is water in the Atlantic Ocean metaphysically implies that there is H₂O in the Atlantic Ocean. But it doesn't epistemically imply it. There are also metaphysically contingent truths that are nevertheless a priori. Consider the truth that if Benjamin Franklin is the actual inventor of bifocals, then Benjamin Franklin is the inventor of bifocals. That is an a priori truth. But it is contingent. The antecedent of the conditional is true in every possible world, while the consequent isn't; thus, there are possible worlds in which the conditional is false. Although 'Benjamin Franklin' and 'the actual inventor of bifocals' are rigid designators, 'the inventor of bifocals' isn't rigid; it can pick out individuals other than Benjamin Franklin in some possible worlds, worlds in which someone else (Newton, say) invented bifocals. The statement that Franklin is the actual inventor of bifocals epistemically implies the statement that Franklin is the inventor of bifocals. But it doesn't metaphysically imply it.

Premise 4 is, I think, false. No microstructural property logically implies a quale. That's because no such property will epistemically imply a quale. But it's another matter whether any property will metaphysically imply a quale.

Stephen Yablo (1992) claims that mental properties are related to certain physical properties of the brain as determinables to determinates. (The claim is essential to his defense of the view that there is mental causation.) In all

of the paradigm cases of determinable-determinate relations it is a priori that whatever has the determinate property has the determinable property. Thus, for instance, maroon is a determinate of red, and so red is a determinable of maroon. It is a priori that whatever is maroon is red. Yablo claims that the case of mental properties and certain physical properties of the brain differs from these paradigm cases in that there are no such a priori connections. He'd thus reject premise 4. He says that, nevertheless, having a brain with a certain physical property is a way of having a mental property; the former metaphysically necessitates the latter. He thus holds that the determinate brain state metaphysically implies the determinable mental state, even though it doesn't epistemically imply it. The claim that there are such metaphysical necessitation relationships between physical properties of the brain and mental properties seems, to me at least, mysterious. This looks to me like mysterious metaphysical necessitation. I reject this sort of physicalist view. (I thus need a different account of how mental causation is possible from the one Yablo offers.)

However, as Kripke showed us, it is a priori that identities are necessary. If every quale Q is such that there is some physical property P such that Q = P, then (given that the relevant physical terms are rigid designators and the qualia terms are too), it is metaphysically necessary that Q = P, even though it will not be a priori. Suppose we found by empirical investigation that the generalization that something has P if and only if it has Q is true, and counterfactual supporting. (Our discovering that is compatible with every theory of phenomenal consciousness we've discussed.) We'd wonder why P gives rise to Q, rather than to some other quale, or no qualia at all? But the question "Why does P give rise to Q?" would be wrong-headed if Q = P. It would be wrong-headed because the presupposition of the question would be false. P doesn't give rise to Q. P *is* Q. If P is indeed Q, then it isn't a fundamental law of nature that something has P if and only if it has Q, or even a law of nature. The generalization is implied by the identity claim.

This view, however, may seem to take us out of the frying pan, but put us into the fire. P will be an objective property. Q will be a subjective property. How could a subjective property be an objective property? How, for that matter, could an objective property be a subjective property? It seems that no subjective property could be an objective property.

Properties, however, must be distinguished from concepts. Properties are ways things might be. Concepts are ways of thinking of things as being or as not being. Two concepts that are not a priori linked can nevertheless answer to the same property. Thus, for instance, the concept of being water is not a priori linked to the concept of being H_2O .²² But the concepts answer to the same property, that is, they apply in virtue of the same property.

The subjective/objective distinction is, in the first instance, an epistemic distinction. It concerns two kinds of ways of understanding, understanding

by empathetically taking up an experiential point of view, and understanding that doesn't require that. What it is for a concept to be subjective is for it to be required for one to fully possess it, to master it, that one knows what it is like to have a certain experience. A concept is objective if there is no such condition on fully possessing it, on mastering it.²³

A property, I claim, is only objective or subjective under a conceptualization, under a concept.²⁴ A property can be subjective under one concept, objective under another. Moreover, one property only epistemically implies another relative to conceptualizations of each. No property under an objective concept epistemically implies any property under a subjective concept. That's why premise 2 is true. No wholly nonqualitative property as such epistemically implies a quale as such.

The view that I favor is that every quale is such that there is some microstructural brain property, some neurobiological property, that is identical with it. The identity will be a posteriori, but necessary. The quale will endow whatever causal powers the neurobiological property endows since it is the neurobiological property. There is one property that we conceptualize in two different ways. This is a view I've taken to calling 'neurobiologicalism.' It's a kind of type physicalism, but not a kind of reductive physicalism. There is a conceptual gap, a gap between subjective concepts and objective concepts, that is, I believe, unbridgeable. It results in an unbridgeable explanatory gap. Qualia *as such* cannot be reductively explained in other terms.

On a separate matter, whether the neurobiological microstructural properties of brains that are qualia are themselves emergent in Broad's sense from other microstructural properties is a question I can leave open. I very much doubt that they will be. But I need take no stand on that issue. Neurobiologicalism is a kind of macrophysicalism as concerns qualia. It doesn't imply microphysicalism.

Of course, a formidable burden that this sort of physicalism faces is to make the case that a subjective concept and an objective concept can apply just in virtue of the same property. It faces other formidable problems as well.²⁵ But all the theories of phenomenal consciousness face formidable burdens. Here I'm just reporting that neurobiologicalism is the horse I'm betting on. There's no mind dust or magic, just an unbridgeable conceptual gap.

Notes

^{1.} Panprotoexperientialism (described in the body of the text) is a view I got at home, growing up. My undergraduate mentor, William L. Stanton, wrote his dissertation at Princeton University. Thomas Nagel was his dissertation supervisor. Stanton is thanked in the final footnote of Nagel (1979). I poured over an early draft of Nagel's paper as an

undergraduate. Over the past thirty-three years, I've included a discussion of panprotoexperientialism (though not, until recently, under that name) in my regularly recurring undergraduate seminar in the philosophy of mind. This is the first paper that I've written focused on it and related doctrines. (Learning about consciousness from a Nagelian perspective, I was frightened away from writing about consciousness until 1999. It just seemed hopelessly hard. Serving as a referee for Chalmers (1996) prompted me to finally dive into the troubled waters and start writing about it. I thank him for the goose. I think.) Tragically, Stanton died in 1984 at age thirty-nine. Without the enormous, selfless attention he showered on me, I would never have had a career in philosophy. I'm so very lucky to have known him. I dedicate this paper to his cherished memory.

- 2. For a critical contemporary discussion of Cartesian-like substance dualism, see Kim (2005, ch.3). For a response, see Audi (2011).
- 3. As Van Cleve (1990) notes, Jonathan Bennett (1984) offers a very similar argument.
- 4. I'll frame the issues here in a way that presupposes three-dimensionalism for objects. But the entire discussion can be recast in a four-dimensionalist framework without affecting any of the main points in this paper.
- 5. This notion of property entailment is due to John Searle (1959).
- 6. Cf. Shoemaker's (2007) notion of a microstructural property. I've helped myself to the notion of an intrinsic property. There is no received view about what it is for a property to be intrinsic. The notion seems intuitive enough for me just to appeal to it here.
- 7. Van Cleve (1990) makes this point.
- 8. I here just use 'proper' in the sense of 'proper part.'
- 9. Chalmers used this apt term in his talk at the conference at which a draft of the present paper and drafts of the other papers in this volume were presented. I should mention that Gregg Rosenberg (2004) uses the term 'panexperientialism' for a novel view of his own. I haven't yet had the opportunity to study it. But it seems to be a kind of panprotoexperientialism rather than panexperientialism properly so-called.
- 10. I'm assuming that the logical inference in question is a priori. For a discussion of the difficult and delicate issue of the a priority of logic, see Field (1996).
- 11. Neutral monism too (at least on the version relevant to the present discussion, there are others) denies premise 2.
- 12. Both Broad (1925) and Samuel Alexander (1920), another leading British emergentist, gave versions of the knowledge argument to argue that qualia are emergent (see McLaughlin 1992).
- 13. Van Cleve (1990) also tries to so explicate it. For a critical discussion of Van Cleve's (1990) claims about Broad's notion of emergence in this context, see McLaughlin (1997/2006).
- 14. Broad (1925) called such laws 'trans-ordinal laws' since, he claimed, they link properties and relations among physical entities at one 'order of reality' with wholes composed of those entities that are at 'a higher order of reality,' a higher order because the wholes possess an ontologically fundamental kind of property not possessed by any of their constituents or by any system composed of those constituents that is organized in any way other than one of the ways that constitute the kind of whole (the kind of system) in question (see McLaughlin 1992).
- 15. Note that exactly the same sort of issue would arise should it turn out that qualia are correlated with functional states (narrow or wide).
- 16. This is reported by Smart (1959). Although Feigl counted the properties in question (emergent properties) as nomological danglers (according to Smart), it has become common in the literature to call the laws themselves (Broad's emergent transordinal laws) nomological danglers.
- 17. Some theorists deny that. They hold 'a collapse view' of quantum mechanics, maintaining that qualia have a role in collapsing the wave function. For a compelling case against that view, see Loewer (2003). I am, to put it mildly, no expert on quantum mechanics. But the experts I know (Tim Maudlin, Shelly Goldstein, Roderick Tumulka, and David Albert) tell me they can make no clear sense of the idea that consciousness collapses the wave function. Arguments by appeal to authority are sometimes apt.
- 18. See the discussion of 'the paradox of phenomenal judgment' in Chalmers (1996).
- 19. See McLaughlin 1989/2008; McLaughlin 1992.

- 20. The notion of alien qualia was introduced in Van Gulick (1993), though not in the context of panexperientialism.
- 21. That's not McGinn's name for the position. That's the name Flanagan (1992) gave it.
- 22. Or so I would argue. See McLaughlin 2007.
- 23. A notion of degrees of subjectivity, and so a corresponding notion of degrees of objectivity can be spelled out using this idea, but I lack the space here to spell it out.
- 24. See Loar 1997; Sturgeon 1994; McLaughlin 2003a. My view differs from Loar's and Sturgeon's in that I spell out the subjective/objective distinction in terms of possession conditions for concepts.
- I try to deal with some of them in McLaughlin (2001; 2003a; 2003b; 2004; 2010; 2012), Hill and McLaughlin (1999), and McLaughlin and Bartlett (2004).

Bibliography

Alexander, Samuel. Space, Time, and Deity. London: Macmillan, 1920.

- Audi, Paul. "Primitive Causal Relations and the Pairing Problem." Ratio 24.1 (2011): 1–16.
- Bennett, Jonathan. A Study of Spinoza's Ethics. Indianapolis: Hackett, 1984.
- Broad, Charlie Dunbar. *The Mind and Its Place in Nature*. London: Routledge and Kegan Paul, 1925.
- Chalmers, David. "Facing up to the Problem of Consciousness." Journal of Consciousness Studies 2.3 (1995): 200-19.
- Chalmers, David. The Conscious Mind. New York: Oxford University Press, 1996.
- Chalmers, David. "Consciousness and Its Place in Nature." The Blackwell Guide to the Philosophy of Mind. Ed. Stephen Stich and Ted Warfield. Oxford: Basil Blackwell, 2002. Repr. in Philosophy of Mind: Classical and Contemporary Readings. Ed. David Chalmers. New York: Oxford University Press, 2002. 247–72.
- Chalmers, David, and Frank Jackson. "Conceptual Analysis and Reductive Explanation." *Philosophical Review* 110.3 (2001): 315–60.
- Coleman, Sam. "Mental Chemistry: Combination for Panpsychists." Dialectica 66.1 (2012): 137–66.
- Dretske, Fred. Naturalizing the Mind. Cambridge: MIT Press, 1995.
- Flanagan, Owen. Consciousness Reconsidered. Cambridge: MIT Press, 1992.
- Field, Hartry. "The A Prioricity of Logic." Proceedings of the Aristotelian Society 96 (1996): 359–76.
- Goodale, Melvyn, and David Milner. "Separate Visual Pathways for Perception and Action." Trends in Neuroscience 15.1 (1992): 20–25.
- Goodale, Melvyn, and David Milner. Sight Unseen. Oxford: Oxford University Press, 2005.
- Haldane, John. "Medieveal and Renaissance Philosophy of Mind." A Companion to Philosophy of Mind. Ed. Samuel Guttenpla. Oxford: Oxford University Press, 1994. 333–37.
- Hill, Christopher. Consciousness. Cambridge: Cambridge University Press, 2009.
- Hill, Christopher, and Brian P. McLaughlin. "There Are Fewer Things than are Dreamt of in Chalmers's Philosophy." Philosophy and Phenomenological Research 59.2 (1999): 445–54.
- Huxley, Thomas. "On the Hypothesis that Animals are Automata, and Its History." *Fortnightly Review* 95 (1874): 555–80.
- Jackson, Frank. "A Priori Physicalism." *Contemporary Debates in Philosophy of Mind*. Ed. Brian P. McLaughlin and Jonathan Cohen. Oxford: Blackwell, 2007. 185–99.
- James, William. The Principles of Psychology. New York: Dover, 1950.
- Kim, Jaegwon. *Physicalism, or Something Near Enough*. Princeton: Princeton University Press, 2005.
- Kripke, Saul. "Identity and Necessity." *Identity and Individuation* Ed. Milton Munitz. New York: New York University Press, 1971. 135–64.
- Kripke, Saul. Naming and Necessity. Cambridge: Harvard University Press, 1980.

- Levine, Joseph. "Materialism and Qualia: The Explanatory Gap." *Pacific Philosophical Quarterly* 64 (1983): 351–61.
- Levine, Joseph. Purple Haze: The Puzzle of Consciousness. Oxford: Oxford University Press, 2001.
- Loar, Brian. "Phenomenal States." *The Nature of Consciousness*. Ed. Ned Block, Owen Flanagan, and Guven Guzeldere, Cambridge: MIT Press, 1997. 597–616.
- Loewer, Barry. "Consciousness and Quantum Theory: Strange Bedfellows." *Consciousness: New Philosophical Perspectives.* Ed. Quentin Smith and Aleksandar Jokic. Oxford: Oxford University Press, 2003.
- Lovejoy, Arthur. "The Meanings of 'Emergence' and Its Modes." *Readings in the Philosophy of Science*. Ed. Philip Wiener. New York: Charles Scribner's Sons, 1927. 585–.
- McGinn, Colin. "Can We Solve the Mind-Body Problem?" Mind 98.391 (1989): 349-66.
- McGinn, Colin. "How not to Solve the Mind-Body Problem." Physicalism and its Discontents. Eds. Carl Gillett and Barry Loewer. Cambridge: Cambridge University Press, 2001. 284–306.
- McLaughlin, Brian P. "The Rise and Fall of British Emergentism." Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism. Ed. Ansgar Beckermann, Hans Flohr, and Jaegwon Kim. Berlin: Walter de Gruyter, 1992. 49–93.
- McLaughlin, Brian P. "In Defense of New Wave Materialism: A Response to Horgan and Tienson." *Physicalism and Its Discontents*. Ed. Carl Gillett, and Barry. Loewer. Cambridge: Cambridge University Press, 2001. 317–28.
- McLaughlin, Brian P. "Color, Consciousness, and Color Consciousness." *Consciousness: New Philosophical Perspectives.* Ed. Quentin Smith, and Aleksandar Jokic. Oxford: Oxford University Press, 2003a.
- McLaughlin, Brian P. "A Naturalist-Phenomenal Realist Response to Block's Harder Problem." Philosophical Issues 13.1 (2003b): 163–204.
- McLaughlin, Brian P., and Gary Bartlett. "Have Nöe and Thompson Cast Doubt on the Neurological Correlates of Consciousness Thesis?" *Journal of Consciousness Studies* 11.1 (2004): 156–67.
- McLaughlin, Brian P. "Emergence and Supervenience." Intellectica 25.2 (1997): 25–43. Repr. in Emergence: Contemporary Readings in Science and Philosophy. Ed. Paul Humphreys, and Mark Bedau Cambridge: MIT Press, 2006. 81–98.
- McLaughlin, Brian P. "On the Limits of *A Priori* Physicalism." *Contemporary Debates in Philosophy of Mind*. Ed. Brian P. McLaughlin, and Jonathan Cohen. Oxford: Blackwell, 2007. 200–24.
- McLaughlin, Brian P. "Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical." *Philosophical Perspectives* 3 (1989): 109–134. Repr. in *Mind and Cognition: A Reader.* Ed. William Lycan and Jesse Prinz. Hoboken: Wiley-Blackwell, 2008.
- McLaughlin, Brian P. "Consciousness, Type Physicalism, and Inference to the Best Explanation." *Philosophical Issues* 20.1 (2010): 266-304.
- McLaughlin, Brian P. "On Justifying Neurobiologicalism." *New Perspectives on Type Identity.* Ed. Simone Gozzano, and Christopher Hill. Cambridge: Cambridge University Press, 2012. 207–29.
- Montero, Barbara. "What Is the Physical?" *The Oxford Handbook of the Philosophy of Mind*. Ed. Brian P. McLaughlin, Ansgar Beckermann, and Sven Walter. Oxford: Oxford University Press, 2009. 173–88.
- Nagel, Thomas. "What Is it Like to Be a Bat?" *Philosophical Review* 83 (1974): 435–50. Repr. in *Mortal Questions*. Ed. Thomas Nagel. Cambridge: Cambridge University Press, 1979. 165–80.
- Nagel, Thomas. "Pansychism." Mortal Questions. Ed. Thomas Nagel. Cambridge: Cambridge University Press, 1979. 181–95.
- Nagel, Thomas. The View from Nowhere. Oxford: Oxford University Press, 1989.

- Nagel, Thomas. "Conceiving the Impossible and the Mind-Body Problem." *Philosophy* 73.285 (1999): 337–52.
- Rosenberg, Gregg. A Place for Consciousness: Probing the Deep Structure of the Natural World. New York: Oxford University Press, 2004.
- Searle, John. "On Determinables and Resemblance, II." *Proceedings of the Aristotelian Society, Supplementary Volumes* 33 (1959): 141–58.
- Shoemaker, Sidney. Physical Realization. Oxford: Oxford University Press, 2007.

Skrbina, David. Panpscyhism in the West. Cambridge: MIT Press, 2005.

Smart, John. "Sensations and Brain Processes." Philosophical Review 68 (1959): 141-56.

Sprigge, Timothy. "Final Causes." Proceedings of the Aristotelian Society 71 (1971): 166–68.

- Sturgeon, Scott. "The Epistemic Basis of Subjectivity." Journal of Philosophy 91.5 (1994): 221-35.
- Tye, Michael. Consciousness, Color, and Content. Representation and Mind. Cambridge, MA: MIT Press, 2000.
- Van Cleve, James. "Mind-Dust or Magic? Pansychism Versus Emergence." *Philosophical Perspectives* 4 (1990): 215–26.
- Van Gulick, Robert. "Understanding the Phenomenal Mind: Are We All Just Armadillos?" Consciousness. Ed. Martin Davis and G. Humphrey. Oxford: Basil Blackwell, 1993.
- Van Inwagen, Peter. "Dualism and Materialism: Athens and Jerusalem?" *Faith and Philosophy* 12.4 (1995): 475–88.

Yablo, Stephen. "Mental Causation." Philosophical Review 101.2 (1992): 245-80.

13

Emergence and Panpsychism

ACHIM STEPHAN

13.1 Introduction

Galen Strawson has posed a dilemma for anyone who is a realist about phenomenal consciousness but rejects panpsychism—the view that every concrete thing has experiences (Strawson 2006). The dilemma is this. Such a person must, then, either embrace reductive accounts of consciousness or else embrace emergence—the view that the experiential emerges from the nonexperiential. But, claims Strawson, any reductive physicalism implies eliminativism about phenomenal consciousness; and emergence is committed to the untenable view that the relationship between the experiential and the nonexperiential is unintelligible, magical. Godehard Brüntrup, for similar reasons, prefers panpsychism to what he calls *inter*-attribute-emergentism—the view that properties of a completely new class of attributes, for example, phenomenal consciousness, emerges abruptly and seamlessly (Brüntrup 2011; Brüntrup this volume). With its upward opacity, he takes inter-attribute-emergentism to be an unreasonably intellectually demanding position. I will argue that both Strawson and Brüntrup overestimate the explanatory resources of panpsychism. In this respect, panpsychism is no better off than emergentism when confronted with the problems philosophy of mind notoriously encounters but is overcharged with metaphysical ballast.

To argue this claim, I will introduce various forms of emergentism and contrast them with varieties of panpsychism, discussing characteristic features such as physical monism, systemic properties, synchronic determination, novelties, and particularly the failure of reductive explanation. It will turn out that—given what we can possibly know about the presumed "mental life" of basic particles—panpsychism does not deliver its promised goods.

13.2 Emergentism Stepwise—Weak Synchronic Emergentism and Diachronic Emergentism

Let us start with emergentism. Three emergentist theories deserve particular attention: weak (synchronic) emergentism, diachronic emergentism, and strong (synchronic) emergentism. Weak emergentism specifies the minimal criteria for emergent properties. Its three basic features—the thesis of physical monism, the thesis of systemic (or collective) properties, and the thesis of synchronic determination—are perfectly compatible with current reductionist approaches. More ambitious theories of emergence have a common base in weak emergentism and can be developed by adding further theses. Diachronic emergentism, for example, acknowledges genuine novelties and discusses their potential predictability; strong emergentism includes the thesis that emergent properties are irreducible. I will examine the more ambitious theories subsequently, but I will not discuss what is known as "radical kind emergence" (van Gulick 2001, 17), a position that rejects the principle of synchronic determination.

Consider weak emergentism first, comparing it with various types of panpsychism (see Brüntrup this volume). The first thesis of current theories of emergence—the thesis of physical monism—concerns the *nature of systems* that have emergent properties. It says that the bearers of emergent features are physical entities only. According to physical monism, all possible candidates for emergent properties such as, for instance, being alive, deciding and acting freely, enjoying the taste of a glass of Rheingauer Riesling, or feeling helpless and lonely are instantiated only by physical systems with a sufficiently complex microstructure. The thesis of physical monism denies there are any supernatural components such as an *entelechy* or a *res cogitans* that are responsible for a system having emergent properties. This means in particular that living, cognizing, and experiencing systems consist of the same basic parts as lifeless or mindless objects of nature.

Physical monism. All entities (whether existing or coming into being) consist solely of physical constituents. Properties, dispositions, behaviors, or structures classified as emergent are instantiated by systems consisting exclusively of physical entities.

In one sense, panpsychism also accepts this picture, since it denies supernatural components. But it differs from emergentism concerning the features of the basic parts. According to emergentism, as introduced and discussed here, basic physical entities are "physicSalist" entities;¹ they have neither a subjective perspective, mental properties such as experiences, nor proto-mental properties such as unexperienced qualities. According to panpsychism and in contrast to emergentism, basic physical entities have—in addition to their physicSalist properties—either proto-mental properties or mental properties together with a subjective perspective.

Physical monism		
Emergentism	Proto-mental property Panpsychism	Mental property panpsychism
Basic physical entities do not have	Basic physical entities have	Basic physical entities have
 a subjective perspective mental properties (e.g., experiences) proto-mental properties (e.g., unexperienced phenomenal qualities) 	 proto-mental properties (e.g., unexperienced phenomenal qualities) they do not have mental properties (<i>e.g.</i>, experiences) a subjective perspective 	 mental properties (e.g., experiences) and a subjective perspective

Whereas the thesis of physical monism places emergent properties and structures within the framework of a physicSalistic naturalism, the second thesis the thesis of systemic properties—delimits the types of properties that are candidates for being emergent. The latter thesis is based on the idea that the general properties of a complex system fall into two classes: a class of properties shared by some of the system's parts, and a class of properties not shared by any of the system's parts. The first class includes, for example, properties such as being extended in a certain way or having a specific weight. The second class includes, for example, breathing, reproducing, feeling itchy, or deciding to retire at fiftynine. These properties are called "systemic" (or "collective") properties.

Systemic properties. A property of a system is systemic if and only if the system possesses it but no part of the system possesses it. Emergent properties are systemic properties.

With respect to systemic properties, emergentism and panpsychism agree in their classification in most respects. For example, both treat being a plane of $4m^2$ or having a weight of 5kg as nonsystemic properties, and both treat properties such as flying, breathing, mating, or feeling itchy as systemic properties. They differ, however, in complex ways with respect to the property of having a subjective perspective and with respect to some mental and proto-mental properties. For all types of panpsychism that ascribe only proto-mental properties to basic physical entities but do not ascribe mental properties to them, having a subjective perspective and/or having mental properties are clear instances of systemic properties. By contrast, for all types of panpsychism that ascribe a subjective perspective all the way down even to basic physical entities, having a subjective perspective is not a systemic property but is instead a fortiori a nonsystemic property. But even forms of panpsychism that ascribe subjective perspectives and other mental states and processes to basic physical entities should allow for some systemic mental properties, for example, feeling itchy, feeling free in one's decision to accept a new job offer, or feeling bored and lonely. To claim the opposite and maintain that basic physical entities could even share those mental properties with us would introduce more magic into our world than any emergentist ever bargained for.

Systemic properties		
Emergentism	Proto-mental property Panpsychism	Mental property panpsychism
Examples of nonsystemic properties: • being a plane of 4m ² • having a weight of Skg	 Examples of nonsystemic properties: being a plane of 4m² having a weight of 5kg having unexperienced phenomenal qualities 	 Examples of nonsystemic properties: being a plane of 4m² having a weight of 5kg having unexperienced phenomenal qualities having a subjective perspective
 Examples of systemic properties: flying, breathing, mating having unexperienced phenomenal qualities having a subjective perspective feeling bored and lonely 	 Examples of systemic properties: flying, breathing, mating having a subjective perspective feeling bored and lonely 	 Examples of systemic properties: flying, breathing, mating feeling bored and lonely

While the thesis of physical monism restricts the type of parts out of which systems with emergent properties may be built, and the thesis of systemic properties characterizes in greater detail the type of properties that might be emergent, the thesis of synchronic determination specifies the type of relationship that holds between a system's microstructure and its emergent properties.

Synchronic determination. A system's properties and its dispositions (to behave in a certain way) depend nomologically on its microstructure. There can be no difference in a system's systemic properties without some difference in the properties or arrangement of its parts.

Anyone who denies the thesis of synchronic determination must either allow for a system to have properties that are not bound to the properties and arrangement of its parts (as radical kind emergentism does) or suppose that other factors, in this case nonnatural ones, are responsible for the fact that systems with identical microstructures can have different dispositions. She would have to admit, for example, that objects may exist that have the same parts in the same arrangement as diamonds but lack diamonds' hardness, that may have hardness 2 instead of hardness 10 on the MOHS-scale. This seems highly implausible. Equally inconceivable is that two microidentical organisms might exist, one of which is viable and the other is not. In the case of mental phenomena, opinions may be more controversial; but one thing seems clear: anyone who believes, for example, that two creatures which are identical in microstructure could be such that one is colorblind while the other can distinguish colors in the ordinary way, does not, in contrast to emergentists and reductive physicalists, subscribe to physical monism.

Like emergentism, panpsychism in its compositional form, seems to accept the idea of synchronic determination. It claims that the "physical level determines all the facts, if 'physical' is taken in the broad sense such that quiddities, that is, the intrinsic natures, are included" (Brüntrup this volume). Panpsychism, however, does not embrace synchronic determination with regard to physicSalist properties. Although it maintains that each duplicate of a specific particle will share both its physicSalist and its intrinsic properties, panpsychism does not claim that particles, which are indistinguishable with respect to their physicSalist properties, are also indistinguishable with regard to their intrinsic properties: the physicSalist properties of an entity do not determine that entity's non-physicSalist properties, its intrinsic nature. Hence, panpsychism allows for the basic components of a complex system to differ in their intrinsic, non-physicSalist properties while their physicSalist properties are the same. This possibility invalidates synchronic determination with respect to the physicSalist properties of a system's components, which are the properties at issue for emergentism.

Thus, one consequence of panpsychism is that it does not offer a robust and accessible metaphysical background, for example, for the search of the neural correlates of consciousness. This result conflicts with Brüntrup's claim that compositional panpsychism just adds nonobservable intrinsic natures to the scientific image, such that these natures "do much of the metaphysical heavy lifting in the philosophy of mind, without getting in the way elsewhere by interfering with the physical laws governing observable physical processes" (Brüntrup this volume). According to panpsychism, two conscious states could differ even if they have identical physicSalist subvenience bases namely, if their subvenience bases differ in their lower-level mental or protomental properties. Since nobody has access either to proto-mental or to mental properties of the basic components of conscious organisms, we have no chance of determining the neural correlates of conscious states.

Synchronic determination		
Emergentism	Proto-mental-property panpsychism	Mental-property panpsychism
There can be no difference in the systemic properties of a system without a difference in the physicSalist properties or in the arrangement of its parts.	There can be no difference in the systemic properties of a system without a difference in its physicSalist properties , its intrinsic properties, or in the arrangement of its parts. There can be a difference in the systemic properties of a system without a difference in the physicSalist properties or the arrangement of its parts, if the intrinsic proto- mental properties of its parts differ.	There can be no difference in the systemic properties of a system without a difference in its physicSalist properties, its intrinsic properties, or in the arrangement of its parts. There can be a difference in the systemic properties of a system without a difference in the physicSalist properties or the arrangement of its parts, if the intrinsic mental properties of its parts differ
		,

(continued)

Emergentism	Proto-mental-property panpsychism	Mental-property panpsychism
Example: Two organisms that	Example: Two organisms that have	Example: Two organisms that have
have identical physicSalist	identical physicSalist microstructures can	identical physicSalist microstructures
microstructures	differ in mental respects	can differ in mental
mental respects	e.g., having a subjective	anxious) if their basic
(e.g., having a subjective	anxious) if their basic parts have different	parts have different intrinsic mental
perspective and/or feeling anxious).	intrinsic protomental properties.	properties.

Synchronic determination

Thus, the two versions of panpsychism differ remarkably from all sorts of emergentism (with the exception of radical kind emergentism), particularly with regard to the properties of a system's basic components and the resulting restrictions to synchronic determinism.

Another difference becomes visible when we turn to diachronic emergentism, a theory of emergence that enriches weak emergentism by taking into account genuine novelties in evolutionary processes (see Stephan 2007, §§ 3.2; 3.6; 4.3).

Novelty. In the course of evolution, exemplifications of genuine novelties occur again and again. Existing building blocks develop new configurations; new structures are formed that constitute new entities with new properties and behaviors.

The thesis of novelty does not by itself turn a weak theory of emergence into a strong one—for reductive physicalism remains compatible with (weak) diachronic emergentism. Yet this thesis implicitly marks the most basic difference between panpsychism and emergentism. While emergentism accepts both socalled *intra*-attribute and *inter*-attribute novelties,² panpsychism disapproves of the latter, that is, of novelties, which are "absolutely new in [their] metaphysical nature" (Brüntrup this volume). Accordingly, for diachronic versions of emergentism, salient examples of genuine novelties include the onset of life, an intra-attribute novelty, and the onset of mind, an inter-attribute novelty (see Alexander 1920, 2: 4–8; Lloyd Morgan 1923, 9–14). Since panpsychism disapproves of inter-attribute novelties, it has to treat both the onset of life and the onset of higher-level mental phenomena as intra-attribute novelties (see. Brüntrup this volume). This move, however, comes at a price. Panpsychism must furnish the world from the very beginning in a way that enables all novelties to be *intra*-attribute. In particular, not even a universe initially devoid of life can have been devoid of mentality. Specific higher-level features of mentality may have evolved but not mentality as such—a claim that is in fact the *mark of panpsychism*: "In the history of the universe there is no clear cut-off point at which experience emerges out of a past that is absolutely void of any experience. Rather, experience was somehow present at the very origin of things, it is a fundamental feature of nature" (Brüntrup this volume).

Genuine novelty		
(Diachronic) Emergentism	Proto-mental-property panpsychism	Mental-property panpsychism
Genuine novelties can be both intra- attribute and inter-attribute novelties.	The only plausible candidates for being genuine novelties are intra-attribute novelties. The occurrence of inter- attribute novelties would imply that something new can evolve from nothing.	The only plausible candidates for being genuine novelties are intra-attribute novelties. The occurrence of inter- attribute novelties would imply that something new can evolve from nothing.
 Examples: the onset of life (an intraattribute novelty) the onset of mentality (an inter-attribute novelty) 	 Examples: the onset of life (an intra-attribute novelty) the onset of higher- level features of mentality (also an intra-attribute novelty) 	 Examples: the onset of life (an intra-attribute novelty) the onset of higher-level features of mentality (also an intra-attribute novelty)

As we have seen, the thesis that genuine novelties occur in the course of evolution does not by itself turn a weak theory of emergence into a strong one; only the addition of the thesis of *unpredictability* will lead to stronger forms of *diachronic* emergentism (see Stephan 1998, 645–47; Stephan 2004, 94– 96). According to the abovementioned basic theses of emergentism, new and possibly unpredictable properties must be instantiated by systems with novel microstructures—otherwise the thesis of synchronic determination would be refuted. Therefore, stronger versions of diachronic emergentism focus on the unpredictability of structures.

The structure of a newly arising system can be unpredictable for various reasons. Its arrangement may be the result (1) of indeterministic processes, (2) of deterministic but chaotic processes, or (3) of an uncompressible unfolding. Discussion among emergentists is mainly limited to the second and third options; only rarely is emergentism concerned with indeterministic processes.

Structure unpredictability. The rise of a novel structure is unpredictable if its formation is either governed by laws of deterministic chaos or is due to an uncompressible unfolding. Any property that is instantiated by the novel structure is likewise unpredictable.

The term "uncompressible unfolding" applies to processes that result in systemic macrostates (or patterns) that can only be derived by *complete* simulations of all interactions at the component's level (see Clark 2001, 116–17; Stephan 2006, 493, 495–96).³ To predict the formation of structures that are governed by deterministic chaos, complete simulations of the underlying microdynamics would be necessary, too. Yet exact simulations are often unavailable for chaotic processes.

In any case, it seems that the feature of structural unpredictability does not mark an important difference between panpsychism and emergentism. Panpsychists agree, for example, that cellular automata (which are paradigms of uncompressible unfolding) are governed by extrinsic properties only (see Brüntrup this volume).

13.3 Emergentism Stepwise—Strong Synchronic Emergentism

Panpsychists' criticism of emergentism is mostly concerned with strong versions of synchronic emergence. Strong synchronic emergentism is of great importance to discussions of the psychophysical problem, particularly for the formulation and analysis of physicSalist but nonreductive positions in the philosophy of mind. Key questions concern the relation between mental and physicSalist properties, for example, whether mental properties, such as having intentional or phenomenal states, can be reductively explained by a physicSalist basis. If we answer "no" and claim—contra panpsychism—that a physicSalist basis is the only basis we can get, we hold a strong emergentist position: we claim mental properties to be irreducible and thus to be emergent in a strong sense. But what does it mean to say that a property cannot be reductively explained? Generally, we ask for reductive explanations when we want to understand why and how a specific entity instantiates a certain property, in fact a property that is only attributed to the system as a whole. The aim of each reductive explanation is to explain (or predict) the dispositions and properties of a system solely by reference to the properties, arrangement and interactions of its components. Hence, to be successful, a reductive explanation must meet several conditions:

- 1. The property to be reduced must be reconstrued (or construed) in terms of its causal or functional role.
- 2. The specified causal or functional role must result from the properties and behaviors of the system's parts and their mutual interactions.
- 3. The behavior of the system's parts must result from the behavior they manifest in systems that are simpler than the system in question.⁴

The third condition, which restricts the reduction base, usually receives little attention in the philosophy of mind. Since mental properties are considered remarkably intractable, those philosophers interested in giving reductive explanations recommend employing anything that could count as a reduction base: the complete physical system including its components; their arrangement, properties, and interactions; as well as all environmental properties relevant to the system's behavior. Thus among others, Jaegwon Kim and Joseph Levine do without the last condition. Both stress, however, that we must first work "into shape" the concept of the property to be reductively explained (Levine 1993, 132). Kim calls this the "priming procedure" in which we must construe, or reconstrue, the property to be reduced relationally or extrinsically (Kim 1998, 98; see Stephan 2004, 96–101).

Because the three conditions for reductive explanations are mutually independent, systemic properties may be irreducible in three different ways:

Irreducibility. A systemic property is irreducible if (1) it is not functionally construable or reconstruable; if (2) the specified functional role does not result from the properties and behaviors of the system's parts and their mutual interactions; or if (3) the specific behavior of the system's components, which microdetermines the systemic property (or behavior), does not result from the behavior of those components in systems that are simpler than the one in question.

Remarkably, and irrespective of all other differences, most philosophers of mind agree that mental properties such as phenomenal and subjective experiences cannot (and will not) be reductively explained in the sense specified by the emergentists and hence should, in this respect, be treated as irreducible. One exception is Michael Pauen (2002), who thinks that future conceptual change might provide a basis even for reductive explanations (but see Stephan 2002). Usually, opinions diverge only on the question whether a failure of reductive explanation implies the irreducibility of mental properties *tout court*. While new-wave materialists claim that the reduction of mental properties might be accomplished by identification with their physicSalist subvenience basis (see, e.g., McLaughlin 2001; Papineau 1998),⁵ panpsychists promise to account for higher-level mentality by enlarging the subvenience basis with either proto-mental properties or lowerlevel mental properties.

In any case, when emergentists claim that mental properties are not subject to reductive explanation, they are not referring to a lack of insight among scientists (in particular neuroscientists) involving a failure to meet conditions 2 or 3; they claim rather that we will fail in principle to meet condition 1, that is, to adequately construct (or reconstruct) phenomenal qualities and subjective experiences in terms of their functional (or causal) role. But if we cannot functionalize a certain mental property, it does not matter how much we know (or might know), say, about the neural correlates of consciousness, or about how various assemblies of neurons behave: that mental property is and will remain irreducible. This was Levine's diagnosis of the so-called explanatory gap some twenty years ago: "What seems to be responsible for the explanatory gap, then, is the fact that our concepts of qualitative character do not represent, at least in terms of their psychological contents, causal roles" (Levine 1993, 134).

Now turn to the panpsychist. From her perspective, the emergentist's position is unintelligible—introducing magic into nature. It makes no sense, or so the panpsychist claims, to assume that we can get something from nothing—experiences from nonexperiential matter (Brüntrup this volume). To beware of this 'metaphysical impertinence,' panpsychism regards nature as being mentally equipped from the very beginning. It assumes that the basic components of beings exhibiting higher-level mentality already have, themselves, some form of proto- or lower-level mentality.

One strand of panpsychism, so-called compositional panpsychism, seems prima facie to try out a reductive route. It claims that the higher-level mental properties of conscious beings result from their organisms' microstructure according to certain compositional principles, thereby assuming that—in addition to their physicSalist properties—the organisms' basic components have intrinsic lower-level mental properties.⁶ But how does the enlarged property set of basic components help us deduce the features of higher-level mentality? Since panpsychism accepts the emergentists' assessment that at least some features of higher-level mental properties cannot be functionalized and therefore cannot be reductively explained, the addition of intrinsic lower-level mental properties to the reduction basis does not improve the situation. Rather, the classical model for reductive explanation would have to be supplemented by an account of those features of higher-level mental properties that cannot be functionalized.

However, the attempt to explicitly account for higher-level mentality in panpsychist terms reveals new obstacles. Among them are well-known problems acknowledged by panpsychists themselves, such as the combination problem (see Brüntrup this volume): Why (and particularly in what arrangement) do intrinsic lower-level mental properties (or minds) yield a higherlevel property (or mind)? How must they be combined to do so? To demonstrate that we are completely ignorant regarding these matters, consider the following intricacies: The body of someone who is falling asleep consists of the very same components in nearly the same arrangement as the very same body a few seconds before when the person was still awake (the same holds true of someone who has just awoken from a coma). So what is it that allows the intrinsic mental properties of the basic components to combine to a higher-level experiential conscious life in one of these states, but hinders them from doing so in the very next moment (or has hindered them in the previous moment)? It can't be the nature of the intrinsic properties belonging to the basic components, since these properties are unchangeable; and it hardly can be the arrangement of the basic components, since they stay in the same place, more or less. Thus, furnishing nature's basic particles with mental features does not solve the problem of conscious experience that the emergentist is said to have. Some features of higher-level mental properties remain irreducible, even if we enhance the reduction basis by adding intrinsic lower-level mental properties.

Another strand of panpsychism, so-called noncompositional panpsychism, accepts this result explicitly. Although it thereby becomes an emergentist position, it insists on a 'decisive' difference between classical (physicSalist) emergentism and panpsychist emergentism. Even if, according to noncompositional panpsychism, higher-level mental properties are irreducible and hence strongly emergent, they are not, as Brüntrup has it, as "superstrongly" emergent as classical emergentism must presume. The reason is that superstrong emergence is "inter-attribute emergence," whereas all other forms of emergence are only "intra-attribute emergence" (Brüntrup this volume). In contrast to superstrong emergentism, panpsychist emergentism need not assume the appearance of something absolutely new in its metaphysical nature. Higher-level mentality, though irreducible, is thought to emerge from arrangements of particles that are endowed with lower-level mentality.

irreducibility		
"Superstrong" physicSalist (i.e., classical) emergentism	Compositional (weak- emergent) panpsychism	Noncompositional (strong-emergent) panpsychism
Some mental properties cannot be functionalized and therefore cannot be reductively explained—they are irreducible. These properties are instantiated by systems consisting of components that altogether lack any mental (or proto- mental) properties.	Although some mental properties cannot be functionalized, they are claimed to be accounted for by the intrinsic nature of their components (plus their extrinsic properties and arrangement). These properties are instantiated by systems that consist of components which themselves have mental (or proto-mental) properties.	Some mental properties are neither functionalizable nor subject to explanation by the intrinsic nature of their components (plus their extrinsic properties and arrangement). These properties are instantiated by systems that consist of components which themselves have mental (or proto-mental) properties.
 Examples: having a subjective perspective having conscious thoughts, ideas, sensations, and feelings 	Examples: • having conscious thoughts, ideas, sensations, and feelings	 Examples: having a subjective perspective having conscious thoughts, ideas, sensations, and feelings

Irreducibility

13.4 Farewell Address

To take stock, higher-level mental properties escape reductive explanations. In this respect, the various strands of panpsychism are no better off than classical emergentism. Whereas emergentism frankly accepts that it has no answer to the problem of how complex organisms consisting of parts void of any experiential features instantiate subjective phenomenal experiences, panpsychism charges the basic entities of nature with proto-mental or lower-level mental properties that we will never become acquainted with (either directly or indirectly). Hence, we will never be able to gain any insight into the (hidden) principles that compositional panpsychism supposes in its account of higher-level mental properties. In particular, we will not understand how an allegedly nonsystemic property such as having a subjective experience might occasionally fail to be instantiated by an organism that normally exhibits it, for example, when it is asleep or in coma.

A noncompositional panpsychism that is ready to accept this outcome subscribes to strong emergentism but ensures itself against emergentism by prescribing a metaphysical pill against magic—it introduces billions of mindful, though lifeless, elementary particles to make up the furniture of the universe. For me, that looks far too costly given the marginal achievements of panpsychism. This is what happens when metaphysics throws parties without inviting science.

Notes

- 1. The term "physicSalism" was coined by Strawson to denote and distinguish standardphysicalism (according to which the "essence of all concrete reality can in principle be fully captured in the terms of *physics*") from "real physicalism," which explicitly accepts that experiential phenomena *are* physical phenomena (Strawson 2006, 4).
- Inter-attribute (mental) novelties are those properties that Charles Dunbar Broad once claimed are bound to their subvenience bases by so-called trans-physical laws (Broad 1925, 79–81).
- 3. Mark Bedau, who introduced the simulation-based notion of emergence in the context of game-of-life environments, dubbed it *weak* emergence to distinguish it from the *strong* synchronic type which comes with irreducibility (Bedau 1997). His weak emergentism, however, is more ambitious than what I called weak and took as the minimal base for all other variants of emergentism.
- 4. An in-depth analysis of these relationships is given in Boogerd et al. (2005).
- 5. New-wave materialists fail, however, to present answers to the reverse task, namely to the question under what conditions, e.g., a newly construed robot has phenomenal experiences. Here it does not help to provide a 'solution' by identification; what we need is an insight into the mechanisms that instantiate experiences; and this is what reductive explanations usually provide (see Stephan 2006, 489–92).
- 6. When I talk of lower-level mental properties, if not stated otherwise, I take this to always include proto-mental properties.

Bibliography

Alexander, Samuel. Space, Time, and Deity. The Gifford Lectures at Glasgow 1916–1918. Vol. 2. London: Macmillan, 1920.

- Bedau, Mark. "Weak Emergence." Philosophical Perspectives: Mind, Causation, and World. Vol. 11. Ed. James Tomberlin, Malden: Blackwell, 1997. 375–99.
- Boogerd Fred, Frank Bruggeman, Robert Richardson, Achim Stephan, and Hans Westerhoff. "Emergence and its Place in Nature: A Case Study of Biochemical Networks." Synthese 145.1 (2005): 131–64.

Broad, Charles Dunbar. The Mind and its Place in Nature. London: Kegan Paul, 1925.

- Brüntrup, Godehard. "Alter Wein in neuen Schläuchen. Die Renaissance des Panpsychismus in der gegenwärtigen Philosophie des Geistes." *Ein Universum voller Geiststaub? Der Panpsychismus in der aktuellen Geist-Gehirn-Debatte.* Ed. Tobias Müller and Heinrich Watzka, Paderborn: mentis, 2011. 23–57.
- Brüntrup, Godehard. "Emergent Panpsychism." *Panpsychism*. Ed. Godehard Brüntrup und Ludwig Jaskolla, Oxford: Oxford University Press, this volume.
- Clark, Andy. *Mindware. An Introduction to the Philosophy of Cognitive Science*. Oxford: Oxford University Press, 2001.
- Kim, Jaegwon. Mind in a Physical World. Cambridge: MIT Press, 1998.
- Levine, Joseph. "On leaving out what it's like." *Consciousness*. Ed. Martin Davies and Glyn Humphreys, Oxford: Blackwell, 1993. 121–36.
- Lloyd Morgan, Conwy. Emergent Evolution. The Gifford Lectures. Delivered in the University of St. Andrews in the Year 1922. New York, London: Williams and Norgate, 1923.
- McLaughlin, Brian. "In Defense of New Wave Materialism: A Response to Horgan and Tienson." *Physicalism and Its Discontents*. Ed. Barry Loewer, Cambridge: Cambridge University Press, 2001. 317–28.
- Papineau, David. "Mind the Gap." Philosophical Perspectives 32.12 (1998): 373-88.
- Pauen, Michael. "Invertierte Schmerzen." Phänomenales Bewußtsein: Rückkehr zur Identitätstheorie? Ed. Michael Pauen und Achim Stephan, Paderborn: mentis, 2002. 268–98.
- Stephan, Achim. "Varieties of Emergence in Artificial and Natural Systems." Zeitschrift für Naturforschung 53c (1998): 639–56.
- Stephan, Achim. "Phänomenaler Pessimismus." Phänomenales Bewußtsein: Rückkehr zur Identitätstheorie? Ed. Michael Pauen und Achim Stephan, Paderborn: mentis, 2002. 342–63.
- Stephan, Achim. "Phenomenal Emergence." Networks 3-4 (2004): 91-102.
- Stephan, Achim. "The Dual Role of 'Emergence' in the Philosophy of Mind and in Cognitive Science." *Synthese* 151.3 (2006): 485–98.
- Stephan, Achim. *Emergenz. Von der Unvorhersagbarkeit zur Selbstorganisation*. Paderborn: mentis, 2007.
- Strawson, Galen. "Realistic Monism. Why Physicalism Entails Panpsychism." Journal of Consciousness Studies 13.10–11 (2006): 3–31.
- Van Gulick, Robert. "Reduction, Emergence and Other Recent Options on the Mind-Body Problem: A Philosophical Overview." *Journal of Consciousness Studies* 8.9–10 (2001): 1–34.

14

Neutral Monism and Panpsychism

LEOPOLD STUBENBERG

14.1 Introduction

Recent work in the metaphysics of consciousness has shown a marked interest in extravagant positions. Panpsychism and neutral monism fall squarely under this heading. Panpsychism has a long history in Western philosophy (see Skrbina 2005) and is the subject of a rapidly growing body of contemporary research. Neither of these claims can be made for neutral monism. Its glory days are tied up with the work of three great philosophers: Ernst Mach (1838– 1916), William James (1842–1910), and Bertrand Russell (1872–1970). But most of the work on neutral monism since that time has been historical. Such modest attention as neutral monism currently enjoys is, primarily, a side effect of the intense work on panpsychism. For the most part, this takes the form of an engagement with Russellian monism—a set of ideas inspired by certain views that Russell defended. This focus on Russellian monism has led to a burgeoning of positions, among them panpsychist interpretations of Russell's neutral monism, as well as neutral monist interpretations of panpsychism.

While the differences between the generic forms of neutral monism and panpsychism are obvious, the current convergence between these two positions makes it hard to differentiate them clearly. But if we go back to the version of neutral monism that Russell himself defended, we shall see that the contrast between it and all the current forms of Russellian monism—be they panpsychist or neutral monist—is quite striking.

So the question arises whether to treat Russell's neutral monism as a mere historical curiosity or as a serious alternative for those who have been driven to embrace panpsychism. I think that a case can be made that Russell's neutral monism should at least count as a live option for those philosophers who are unafraid to take seriously unpopular views.

14.2 Neutral Monism Introduced

Neutral monism is a doctrine about the nature of the ultimate constituents of the world and their relation to the rest of what exists. The monism consists in the fact that the ultimates are all of one kind. The neutrality consists in the fact that entities belonging to this kind are neither mental nor material and, in that sense, neutral between these two poles. And all nonultimate entities, be they material or mental, are viewed as reducible to or constructible from the neutral ultimates.

According to this view, your current thoughts and feelings, and the brain in which they are presumably occurring, are, ultimately, made of the same material: neutral elements that themselves exemplify neither the mental features of your thoughts and feelings, not the physical features of your brain. Mind and matter, material and mental features are therefore not part of the ultimate furniture of the world. In the last analysis they resolve themselves into neutral ultimates.

Physicalistic reductionism tries to reduce mind to matter; idealistic/mentalistic reductionism tries to reduce matter to mind; neutral monism tries to reduce mind and matter to some third, neutral kind of thing. The task facing the neutral monists is huge. First, they must specify the nature of, and secure the existence of, the neutral realm. And then they face a most daunting reductive challenge: they have to show, at least in principle, how everything we know—every mental and physical entity—can be reduced to or constructed from this new and as yet unknown basis.

14.3 The Neutral Entities

The first challenge facing neutral monism is to identify the neutral elements. Something is neutral, in the relevant sense, just in case it is neither mental nor physical. Therefore the shape of one's neutral monism will be determined by the way in which one draws the mental/physical distinction. Were one to hold that the mental and the physical jointly exhaust all there is, then neutral monism would be ruled out: in this setup there is no room for a neutral realm of entities that are neither mental nor physical. But most ways of drawing the mind-body distinction will leave room for neutral entities.

Though there have been few proposals as to the sorts of entities that might qualify as neutral, they have ranged across quite disparate areas. Kenneth Sayre, for example, has argued that information (in the strict Shannon/Weaver sense) is the sort of entity that can serve as the basis of a neutral monist theory (Sayre 1976). The idea that abstract entities might serve as the neutral basis has found an early defender in Edwin B. Holt, who held that the fundamental neutral elements are logical and mathematical entities (Holt 1914, especially 154–60). And David Chalmers's speculations about the place of information in the world may also belong into this camp (Chalmers 1996, ch. 8). While the claim that abstract entities are neutral is very plausible, the suggestion that concrete objects might be reduced to or constructed from abstract objects is deeply problematical and has found few adherents.

As we shall see in more detail later, the famous neutral monists—Mach, James, and Russell—sought the neutral basis in a different direction. As a first approximation, we can say that they took the sensory qualities given in experience to be the neutral elements from which all else was constructed. But most who considered this suggestion rejected it on two grounds: the proposed neutral entities seemed mental rather than neutral; second, the prospects of constructing physical and, to a lesser degree, mental entities from this slender basis seemed little better than building concrete objects out of abstract entities.

14.4 Panpsychism Introduced

"Panpsychism is the view that all things have a mind or a mind-like quality" (Skrbina 2014). The current debate frequently concerns a somewhat narrower doctrine that focuses on physically fundamental entities and experiential qualities. Here, for example, is David Chalmers's characterization of the doctrine: "I will understand panpsychism as the thesis that some fundamental physical entities are conscious: that is, that there is something it is like to be a quark or a photon or a member of some other fundamental physical type" (Chalmers 2015, 246–47).¹

The ultimate constituents of the world (or at least some of them) are "fundamentally physical *and* fundamentally mental" (Chalmers 2015, 248).² And their mentality consists in their having experiences or being conscious, that is, in it's being like something to be this quark, in the quark's having a point of view, in it's being a subject of experience—a microsubject with microexperiences—but an experiencer nevertheless.

On the panpsychist picture, mind and matter, material and mental features, are part of the ultimate furniture of the world. Neither is reduced to the other; and, unlike in the case of neutral monism, there is no neutral, third nature in play. Panpsychism can rely on physics to specify the ultimates and to explain how the physical ultimates are related to the other physical entities. But in postulating microexperiences with their corresponding microselves, panpsychism incurs the obligation to tell us what these things are, how they are related to each other and to the ultimates provided by physics, and how the
microsubjects and their microexperiences are related to the macroversions of themselves—the selves we are and the experiences we have.

14.5 Neutral Monism and Panpsychism Are Incompatible

These brief sketches of neutral monism and panpsychism suffice to highlight the principal difference between them. According to panpsychism, both mind and matter are coequal and fundamental features of the world. According to neutral monism neither mind nor matter are fundamental features of the world; both are constructions from (or are reducible to) a third kind of thing that is more primitive than either of them. The disagreement about the ultimate building blocks of the world could not be more pronounced.

And this difference is reflected in the way in which these two theories envision the relationship between the ultimate building blocks of the world and the world of matter and mind as we know it. In the panpsychist world, mind and matter as we know them in the macroworld, arise from the combination or aggregation of microversions of themselves. The 'size' of the phenomena increases; but the basic characteristics of mentality and materiality trace all the way back to the ultimate building blocks of the world. Microphysical and micromental properties can combine to form new physical and mental properties—but these are resultant properties, properties that arise through addition and subtraction of properties of the same basic kind.³ Emergence, be it radical or effected by some process of construction, is ruled out.⁴

The neutral monist will agree with the panpsychist that radical emergence is off the table; but in other respects the disagreement is complete. All mental and all material phenomena, no matter what their size may be, are ultimately reducible to (constructible from) phenomena that are themselves devoid of these features. Everything bottoms out in neutral elements. Mind and matter are real features of the neutral monist world, but their reality consists in the fact that they are, in Fodor's memorable words, "really something else" (Fodor 1987, 97).

Given only the preceding schematic sketches of neutral monism and panpsychism, it is difficult to motivate a choice between them. A philosopher committed to either one of these views, will find the other one unacceptable. Most philosophers, who have considered them at all, find them both unacceptable. The small group of philosophers for whom these theories are live options are united by their rejection of all the mainline 'solutions' to the mind-body problem. Those philosophers will share many convictions, for example: the rejection of physicalistic reductionism; the rejection of radical emergence; the rejection of Cartesian dualism, the conviction that consciousness is real. But there remains room for disagreement, the most important of which will concern one's stance on the question whether the mind is reducible *at all*. If reduction (or construction) is ruled out altogether, neutral monism is unacceptable. If the reduction (or construction) of the mental is still on the table, the choice between the two theories will have to await their more detailed articulation.

14.6 Neutral Monism Collapses into Panpsychism

The preceding sketches create the impression that the differences between neutral monism and panpsychism are obvious and important. But a look at some of the most interesting recent literature on panpsychism raises doubts about this simple picture. The idea of Russellian monism plays a central role in this line of thought.

Russellian monism grows out of a cluster of ideas that Russell defended: (i) physics is solely concerned with structural/relational/extrinsic/dispositional properties; (ii) such properties are grounded in nonstructural/nonrelational/ intrinsic/categorical properties; (iii) experience familiarizes us with a class of properties that can play this grounding role (Alter and Nagasawa 2012, 70– 71). Russellian monists accept (i)—the thesis of structuralism about physics—and (ii)—the thesis of quidditism: the view that structural properties are grounded in nonstructural properties (see Chalmers 2015, 254).⁵ But there is considerable debate about (iii)—concerning the question of the nature of the quiddities.

A number of panpsychists saw a natural affinity between Russellian monism and panpsychism. If the quiddities are taken to be phenomenal properties, Russellian monism turns into panpsychism—into the view that Chalmers calls Russellian panpsychism: "On this view microphenomenal properties serve as quiddities and also serve as the grounds for macrophenomenal properties. That is, microexperience constitutes macroexperience while also playing microphysical roles" (Chalmers 2015, 255). Occasionally this elegant way of integrating the phenomenal and the physical is attributed to Russell himself, thus turning a paradigmatic neutral monist into a panpsychist (see Coleman 2009, 87; Alter and Nagasawa 2012, 77). If one can show that Russell's neutral monism—the most detailed version of neutral monism on record—collapses into panpsychism, then one has gone some distance toward establishing that we are not dealing with two radically distinct doctrines. The fact that the versions of neutral monism advocated by Mach and James are sufficiently similar to Russell's to be open to the same interpretation further strengthens this argument. While I believe that the panpsychist interpretation of Russell should be resisted, it has to be said that the suspicion that the standard version of neutral monism (Mach, James, Russell) is nothing but a thinly cloaked version of panpsychism has accompanied neutral monism from its inception (see Stubenberg 2014).

14.7 Panpsychism Collapses into Neutral Monism

But we also find the opposite movement in the literature: the suggestion that panpsychism might best be understood as neutral monism. Chalmers (2016), in whose hands Russellianism unfolds into an astonishing variety of different forms, takes us on a journey that leads from Russellian panpsychism to Russellian panqualityism.⁶ Chalmers finds flaws in all of the many forms of Russellianism he investigates, including the final, panqualityist version. But what is important for us to note at this point, is that his final version of the view is a form of neutral monism.⁷ On this view the physical ultimates *instantiate but do not experience* the sorts of qualities that are present to us when, for example, we experience redness. The ultimates have the quality redness, but they do not, in Chalmers's terms, have the phenomenal property redness. That is, they do not represent redness; they are not aware of redness; they are not acquainted with redness. They simply are red. And to be that, they need not be a subject. In this way, the move from phenomenal properties to simple qualities transforms Russellianism from panpsychism to neutral monism.

One can question whether Russellian panqualityism should count as a version of neutral monism. The position is Russellian in that it takes these simple qualities to function as quiddities, grounding both physical and mental properties. The position is monistic because "the world on this view consists in quiddities connected by laws of nature" (Chalmers 2015, 262). But one may wonder whether the case for neutrality is equally strong: have we been shown that the simple qualities that serve as quiddities are neither mental nor physical? The case for the nonmentality of the simple qualities rests, presumably, on the claim that these qualities are distinct from phenomenal properties—they are not qualities of experiences, nor do they involve subjects of experience. But what makes them nonphysical? The first point one can make is this: the simple qualities are not the structural/dispositional/relational/extrinsic properties that physics deals with, that is, they are not narrowly physical. But if one also entertains a notion of broadly physical properties—comprised of the narrowly physical properties plus the properties that realize them (Chalmers 2015)—then simple qualities qualify as physical in this broader sense. This does raise the question whether we are dealing with a version of physical rather than neutral monism.⁸

Though not conducting his discussion within the framework of Russellianism, Sam Coleman has reached a strikingly similar conclusion: "Panpsychists should be neutral monists" (Coleman 2014, 1). Unable to answer the question of how the many microsubjects of the ultimates that constitute you can form the macrosubject you are, the panpsychist must banish the microsubjects and hold that "the ultimates possess phenomenal qualities, despite not being subjects of those qualities" (Coleman 2014, 2). Phenomenal properties, or qualia, "must be divorced from subjectivity—the awareness of qualia by subjects" (Coleman 2014, 21). And since macrosubjects cannot magically emerge out of a sea of subjectless physical ultimates, "subjecthood [the existence of macrosubjects] must be susceptible of a reductive treatment" (Coleman 2014, 21). According to Coleman, this is tantamount to neutral monism: "A position taking qualities as fundamental features of matter, but which makes subjective awareness of qualities a relational (thus reducible) affair, is *neutral monism*, of the sort that James later inclines towards" (Coleman 2014, 21).

Here, too, we can ask whether the resulting position really lives up to its neutral monist billing. Have we been given reason to believe that the qualities are neutral? A little while ago we had occasion to wonder whether we are dealing with a neutral or a physical monism; now I want to raise the question whether we have reason to think that we are dealing with a mental monism. The idea that the qualitative is a paradigm of the mental is not new. In earlier work Coleman made a forceful case for just this view (Coleman 2009, 94-106). The claim he sets out to defend is that "sufficient consideration of the notion of the intrinsically qualitative shows it to be indistinguishable from the notion of the qualitatively experiential. The idea of that which has an absolutely intrinsic way of being just is the idea of the conscious-experiential" (Coleman 2009, 94). It is on this basis that we "should put consciousness at the heart of ontology, and embrace panexperientialism" (Coleman 2009, 94). On this view, it is the qualitative itself (micro or macro), and not its relation to a subject, that secures its mental, that is, nonneutral, nature. If there is merit to this view, the journey from panpsychism to neutral monism is more difficult than it may appear.

14.8 Panpsychist Neutral Monism?

Is there a much quicker path to the conclusion that neutral monism and panpsychism are compatible? David Skrbina has argued that panpsychism is a metatheory, "a theory about theories, a framework, which says: However mind is to be conceived, it applies, in some sense, to all things" (Skrbina 2014; 2005, 2-3; see also Clarke 2004, 9). Accordingly, consideration of some of the more popular theories of mind yields the following list of panpsychist versions of these theories: panpsychist substance dualism, panpsychist functionalism, panpsychist identism, and panpsychist reductive materialism (see Skrbina 2014; 2005, 2-3). Panpsychist neutral monism does not show up in this list. But perhaps it should. The existence of materialist versions of panpsychism in functionalist, identist, or reductive forms-suggests that there might be analogous neutral monist versions of panpsychism. Why not have a panpsychism in which mental properties are reduced to, or identified with, or functionally realized by neutral (rather than physical) properties? The viability of any one of these proposals would suffice to show that neutral monism and panpsychism are compatible. On the face of it, the panpsychist versions of neutral monism may seem no more far-fetched than the panpsychist versions of functionalist, identist, or reductionist materialism. In that case, the suggestion of some deep tension between neutral monism and panpsychism would seem misguided—simply a reflection of an insufficiently broad conception of what panpsychism really is.

But I think that we should resist the quick path to the compatibility of neutral monism and panpsychism. For the proposal to countenance functionalist, identist, or reductive materialist versions of panpsychism turns on rejecting a premise of one of the main arguments for panpsychism—the so-called genetic argument for panpsychism. Thomas Nagel, for example, formulates this premise as follows:

Nonreductionism. Ordinary mental states like thought, feeling, emotion, sensation, or desire are not physical properties of the organism behavioral, physiological, or otherwise—and they are not implied by physical properties alone. (Nagel 1979, 181)

The failure of physicalistic reductionism—broadly construed as the failure of any form of property reduction or identification—combined with the rejection of brute emergence are the two main insights that drive the current revival of interest in panpsychism (see e.g., Nagel 1979; Strawson 2006). And if we take this as a reason to reject the (broadly) materialist versions of panpsychism that Skrbina envisions, we should also refrain from endorsing a neutral monist version of panpsychism. For although the principle of nonreductionism, as stated by Nagel, does not explicitly rule out the possibility that the mental be reduced to or identified with the neutral, the principle is clearly meant to extend to this case as well. And while it may be true that panpsychism is not committed to the principle of nonreductionism and that the genetic argument is not an essential stepping-stone on the route to panpsychism, it seems that cost of this way of resisting the foregoing criticism is considerable: it forces the panpsychist to give up one of the premises in what is, perhaps, the strongest argument in favor of panpsychism.

But it may seem that the preceding thought cannot be right. Take Galen Strawson's position, for example. He relies on the genetic argument to advance a panpsychist version of materialism: "Experience is 'really just neurons firing'" (Strawson 2006, 7). That sounds like straightforward reductionism. Does this mean that Strawson denies the nonreductionist premise on which his argument relies? Not quite. He does not reduce experiential properties to the properties currently found in physics, nor does he identify them with such. Instead he holds that the experiential properties, in all of their qualitative glory, just are physical properties—physical properties *in addition* to those that play a role in current physics. And so long as one does not define physical properties as being nonmental, this is a viable strategy.⁹

Can one employ a similar strategy to arrive at a neutral monist version of panpsychism—can we simply declare that mental properties just are additional neutral properties? The answer is no. While it may be true that the physical is not, by definition, nonmental—thus allowing us to categorize the purely mental as physical—things are different with respect to the neutral. The neutral is defined as being neither mental nor physical. So the proposal that mental properties, taken just as such, simply are additional neutral properties, amounts to the untenable view that there are properties that are both mental and nonmental.

14.9 The Ecumenical Solution

While I do not think that the tension between panpsychism and neutral monism can be defused by noting the metatheoretical character of panpsychism, the considerations that seemed to show that panpsychism and neutral monism collapse into each other do carry some weight. It seems that we can conclude at least this much: in the current discussion the concepts of panpsychism and neutral monism have become considerably more fluid. A sharp demarcation between the two doctrines may no longer be possible. Perhaps the initial attempt to set them up as clear and irreconcilable rivals is misguided. And perhaps the correct answer to the question: 'Which should we choose, neutral monism or panpsychism?' is that we should not choose one over the other, but that we should work toward a view that includes what is best in both of them.

14.10 Russell's Neutral Monism

But rather than close on this conciliatory note, I want to look at Russell's version of neutral monism—the most worked-out version of neutral monism on record—and see how it relates to panpsychism. I will argue that Russell's own story is very different from the various Russellianisms that are currently discussed in the literature. And while it is possible to reinterpret Russell's story in idealist (see Bolender 2001), or panpsychist (see Stubenberg 2015), or physicalist (see Landini 2011; Wishon 2015) ways, I hold that Russell's own version of neutral monism is (or at least strives to be) opposed to any kind of view that makes mind, or matter, or both fundamental.

14.11 Neutrality—Events and Percepts

Russell's neutral monism grows out of and replaces his earlier sense-datum theory. He used to hold the dualist view that in perception the mind (nonphysical) acquaints itself with a sense-datum. Initially Russell maintains that sense-data are nonmental (Russell 1912); in his later writing on sense-data he adds that they are physical: "I hold strongly that the sense-datum is not mental—indeed my whole philosophy of physics rests upon the view that the sense-datum is purely physical" (Russell 1915b, 391). The nonmental character of sense-data strikes Russell as obvious: what is seen, or more generally, what is sensed, has none of the mentality that is characteristic of "occurrences which are indubitably mental: believing, doubting, wishing, willing, being pleased or pained . . . seeing, hearing, smelling, perceiving generally" (Russell 1915a, 130). This, he seems to think, is a matter that is clear on reflection: "It is not difficult to see that that colours and noises are not mental in the sense of having that intrinsic peculiarity which belongs to beliefs and wishes and volitions, but not to the physical world" (Russell 1915a, 132). The additional step that leads him to say that the sense-data are physical is grounded in his changing thoughts about the nature of matter. So long as matter is taken to be a permanent substance, the ever-changing sense-data cannot count as material nor can they function as constituents of matter. But once he sees that "matter as it appears to common sense, and as it has until recently appeared in physics, must be given up" (Russell 1927b, 164) the way is free to classify sense-data as physical:

So long as it is supposed that the physical world is composed of stable and more or less permanent constituents, the fact that what we see is changed by changes in our body appears to afford reason for regarding what we see as not an ultimate constituent of matter. But if it is recognised that the ultimate constituents of matter are as circumscribed in duration as in spatial extent, the whole of this difficulty vanishes. (Russell 1915a, 134)

His subsequent (and absolutely pivotal) insight that the self/mind, together with its mental acts, is undiscoverable, leads him to abandon this dualistic theory and embrace a nonrelational, monistic theory of perception instead. What he used to conceive as a complex relational fact—a mind related to an object via an act—now strikes as a simple datum: "The sensation that we have when we see a patch of colour simply is that patch of colour ... the patch of colour and our sensation in seeing it are identical" (Russell 1921, 142-23). In his mature work on neutral monism, Russell calls the resulting, unstructured occurrences—these patches of color (as well as sounds, smells, tastes, etc.)—percepts. And percepts are neutral. The percepts inherit their nonmentality from the sense-data which they replace.¹⁰ But what makes the percepts (unlike the sense-data that they replace) nonphysical? According to Russell, percepts are paradigms of events—of entities "having a small finite duration and a small finite extension in space ... something occupying a small finite amount of space-time" (Russell 1927b, 287). Events are the chosen building blocks of his neutral monist world, and percepts are the events with which experience familiarizes us. So, for example, "hearing a tyre burst, or smelling a rotten egg, or feeling the coldness of a frog" (Russell 1927b, 287) are typical events that we know from experience. And if matter is going to be constructed out of events, then the events themselves will not count as material: "Matter and motion ... are logical constructions using events as their material, and events are therefore something quite different from matter in motion" (Russell 1927b, 292).

Russell does not limit the construction materials to the events that happen to be percepts; he allows inferences to a multitude of events that are not percepts. The totality of these events comprise the building materials from which to construct all physical and mental entities.

14.12 The Method of Construction

How is one to conceive of a construction out of percepts (and similar inferred events)? How does such a construction proceed? The following example gives us a sense of the mode of construction Russell has in mind:

If we ... confine ourselves to the world of one man's experience, we can easily define an 'instant' in his life. It will be a group of events, all belonging to his experience, and having the following two properties: (1) any two of the events overlap; (2) no event outside the group overlaps with every member of the group. (Russell 1927b, 288)

And he generalizes this model to the case of point-instances in space-time:

By a slightly more complicated but essentially similar method, we can define a point-instant in space-time as a group of events having two properties analogous to those used just now in an 'instant' in one biography. (Russell 1927b, 288–89)

What he proposes is that the construction of point-like physical particles can proceed along similar, if somewhat more complex, lines (Russell 1927a, ch. 28). The resulting picture is that of a world filled with overlapping events. The behavior of some groups of overlapping events is described by the laws of physics; these groups we call electrons, protons, and so forth. In Russell's world, events—not substances, or things, or particles—are the fundamental reality. "It is a mere linguistic convenience to regard a group of events as states of a 'thing,' or 'substance,' or 'piece of matter'" (Russell 1927a, 284). It follows that "electrons and protons . . . are not the stuff of the physical world" (Russell 1927a, 386). More generally, "bits of matter are not among the bricks out of which the world is built. The bricks are events, and bits of matter are portions of the structure to which we find it convenient to give separate attention" (Russell 1924/1956, 329).

14.13 Structuralism and Russell's Use of It

Russell's structuralism—the thesis that physics tells us about the structure, but not the intrinsic nature of things—features prominently in contemporary Russellian thought. But the lesson that Russell draws from structuralism is quite different from the lesson drawn by the Russellians. The latter seize on the idea that structural/relational/dispositional properties need some sort of grounding: no relations without relata; no dispositions without categorical bases; no structure without something to be structured. And in their search for entities that can play this grounding role, they again let themselves be inspired by Russell: his percepts—events whose intrinsic nature is known to us directly—are the paradigms for their quiddities or inscrutables that do the grounding work.

Russell's way of bringing structuralism and percepts together is quite different. The structuralism of physics gives us free hand to construct the entities of physics from whatever materials and in whatever way we see fit. The only constraint on our constructive activity is that the resulting entities have the structural properties that physics ascribes to them.¹¹ The role the percepts (and all the other events inferred from them) play is not to ground structural, relational, dispositional, etc. properties. Their role is to serve as a supply of entities—entities that are better known and therefore less questionable than those postulated in physics-that can be arranged into groups that exhibit the structural properties that physics talks about. This is a straightforward application of Russell's supreme maxim of scientific philosophizing: "Wherever possible, substitute constructions out of known entities for inference to unknown entities" (Russell 1924/1956, 326). The point of this maneuver is both ontological-we avoid the commitment to the existence of the entities that have been replaced by constructions-and epistemological-we no longer incur the risk of error that is inevitably associated with an inference to the existence of an entity that is not directly given.

Structuralist ideas show up a second time in Russell's thinking. The percepts are the only events we know directly. That is, we know them not merely abstractly, but by grasping their intrinsic qualitative nature—a color, a smell, a taste, etc. But all other events, all the events that are not data for us, are known only inferentially. And Russell holds that while these inferences do tell us something about the structural features of the inferred events, they are silent about their intrinsic characteristics. That means that we know nothing about the intrinsic character of the vast majority of events that go into the construction of the vast majority of physical entities. Only in the case of the construction of the physical particles that make up our own brain do matters stand differently: in this special case we have immediate knowledge of some of the events that figure in the construction of some of the particles of our brains.

14.14 Russell's Views Versus Current Russellianisms

All this makes for a picture that is quite strikingly different from the various forms of Russellianism we saw earlier. For the sake of concreteness, let's focus on a particular ultimate—an electron, say—in your brain. Some of the visual, auditory, gustatory, and so on percepts you are currently enjoying may be among the group of events that go into the construction of this electron. Let's say you are sipping a beer and savoring your gustatory beer percept. Assume that this beer percept goes into the construction of the electron—that is, it

overlaps in certain complex ways with lots of other events occurring in your brain. Does this fact give us any reason to maintain that the electron must be enjoying the beer percept? It does not seem so. Nothing about the group of events (including the percepts) that plays the role of the electron suggests that this group should be treated as experiencing the qualities of the events that compose it or as experiencing anything at all. The reasonable assumption is that it is like nothing at all to be this electron; it is not a subject of experience; it has no point of view. And this would be so even if it turned out that it is wholly composed by the percepts you are currently having. So, in this regard, Russell's story is utterly unlike that panpsychist story: no microsubjects and no microexperiences for the fundamental physical entities.

Is it, then, any closer to the panqualityist story? Not really. The reality of the constructed electron consists in the fact that a number of events, among them the beer percept, are grouped in a certain way and that this group evolves according to certain physical laws. The beer percept is in this group; but there is not, in addition to this group, and extra thing—the electron—that has or exemplifies or instantiates the property that composes the event of tasting the beer. Perhaps this is a useful analogy: a grain in the heap of sand is perfectly round. Does that make the heap itself be perfectly round? No, of course not. So the panqualityist suggestion that the ultimates have/exemplify/instantiate but do not experience the qualities in question brings us no closer to Russell's own view. On Russell's view the fundamental physical entities are not clad in microversions of the qualities we know in experience. The only qualities known to be involved are the macroqualities that compose some of the events (the percepts) that go into the construction of the fundamental physical particles of your brain.

Panpsychism and panqualityism are agreed in taking the electron as an ultimate, as a fundamental part of their ontologies. Their disagreement concerns the manner in which experience or experienced qualities inhere in this fundamental particle: Does the electron have qualitative experiences? Or does it exemplify experiential qualities (without experiencing them)? By agreeing to assign a central role to the electron, and much struck by its extreme smallness, both parties are lead to countenance microqualities—smaller versions of the qualities we know from our own experience—that are supposed to be either experienced by microsubjects, or to be exemplified by (self-less) microparticles. But with microqualities and microsubjects come interesting questions about how these are related to their macroversions. This is the source of the many combination problems that currently plague panpsychism and panqualityism (see especially Chalmers 2015 and this volume).

Russell's approach does not invite the postulation of microqualities or microsubjects. Of course Russell countenances electrons and other microparticles.

But they are not ultimate or fundamental in the grand scheme of things. Only the neutral elements—percepts and similar inferred events—qualify for this role. And once the electron is seen to be a mere construction, it is no longer tempting to see it as a subject of experience or as a bearer of special qualities. There is, therefore, no need to introduce electron-sized microqualities or micropercepts into the system. When viewed from this perspective, the attempt to integrate the realms of mind and matter by somehow 'attaching' small amounts of mentality to small amounts of matter, seems quite wrongheaded. The abstractness of physics, together with Russell's method of construction, open up the possibility of a completely different way of bringing mind and matter together. Events—among them our own, full-sized percepts—are used to construct entities that live up to the structural specifications laid down by physics. The percepts come first; the physical particles are, logically speaking, late arrivals on the scene. Our own percepts are all the percepts involved, and they find a natural home as members of the groups of events that are the physical particles that form our brains. This story does not prove that the electron neither has percepts nor exemplifies the sorts of qualities whose instantiation composes percepts. But nothing in this story so much as suggests that anything as strange as this might be the case.

Questions about how small subjects, small experiences, small experiential qualities can come together and form their large counterparts do not arise. There are only your (large) percepts, composed of (large) qualities you experience. They are among the fundamental elements of Russell's neutral monism every bit as fundamental as all the other (inferred) events, whose intrinsic nature—being entirely unknown—may or may not be similar to the intrinsic nature of our percepts. On this view the experiential does not magically emerge from the nonexperiential; nor is the experiential reduced to something that it is not. The experiential, in the form of the having of percepts, enters the world at the fundamental level. If this is right, this view accomplishes much that the panpsychist strives for.

14.15 Problems and the Quality/Awareness Gap

Of course there are many problems. Some have to do with how well we can understand Russell's exposition of neutral monism. Perhaps the main question here concerns the alleged neutrality of the basic elements of neutral monism—events in general and percepts in particular.¹² General questions about the method of construction and about the details of the construction of particular mental and physical entities may be felt to be no less urgent.¹³ But even if we give Russell the benefit of the doubt, and we assume that the theory can be made sense of, we may wonder whether his neutral monism really does deliver what it promises. In particular I want to raise the question whether Russell's neutral monism succumbs to the same problem that Chalmers sees as the nemesis of Russellian panqualityism.

According to Russellian panqualityism, qualities such as "redness, greenness, heat and so on," but more likely qualities "more austere than this, but . . . at least similar in kind" (Chalmers 2015, 272) are quiddities: they serve as the categorical bases for the dispositional properties that physics ascribes to these particles; and they are the grounds of your experience. Now Chalmers wonders whether it could not be the case that all of these qualities are in place, yet nobody is aware of any of them. In that case you remain unconscious—all the quiddities are in place, yet you have no experience, nothing is like anything for you. This is the problem raised by the quality/awareness gap. Does it also arise in Russell's neutral monism?

I do not think so. The only features that Russell ascribes to the microparticles of your brain are the structural properties of physics. He has nothing to say about whether these particles have (instantiate, exemplify) any qualities (either micro or macro; either full-blown or austere), nor does he think that such qualities are quiddities. Of course he does talk about qualities—the qualities instances of which compose the events that are your percepts. The qualities composing your percepts do enter the microparticles of your brain but only as members of the overlapping groups of events that jointly construct those particles. These qualities are not quiddities: they do not have the job of serving as categorical bases of the properties that physics ascribes to the particles they constitute; and they do not have the job of grounding macroqualities. They are themselves macroqualities-qualities of which you are aware because instances of those qualities compose the events that are your percepts.¹⁴ To have a red-percept is like something for you—it is to see red. In this scheme, the question whether you could be unconscious while all of your particles, together with their quiddities are in place, does not arise. There is no place for quiddities in this system. The qualities that are involved play a very different role.

My impression is, then, that the structural differences between Russellian panqualityism and Russell's neutral monism are sufficiently large to keep the quality/awareness gap—in the form in which Chalmers posed it—from arising in Russell's neutral monism. But perhaps the same sort of issue does arise in neutral monism, albeit in a slightly different way. The worry that immediately comes to mind is this: does a story like Russell's—basically a bundletheoretic story—have any resources to explain how events get bundled in the way they do, and to explain what it is about certain bundles that makes them into subjects of experience. More specifically, can Russell explain what it is that makes an event (that is a part of the bundle you are) into one of your percepts? What sorts of relations among events make a given event be a percept? These relations must bring it about that the quality that composes this event is revealed to you in form of an experience you have. Are there any relations that can accomplish this feat? And once all this is before us, we then need to ask: Is it possible that there be all these events, bundled just so, and interrelated just so, while the person who is this bundle does not have any experience at all? If the answer is yes, then Russell's neutral monism founders on the same rock as Russellian panqualityism.

14.16 Conclusion

If one insists on traditional textbook definitions of neutral monism and panpsychism, the two doctrines are incompatible. The current debate is far removed from this simplistic opposition. We are witnessing a great unfolding of possible positions, making the radical opposition between the two views appear increasingly artificial. But if we turn back to Russell's neutral monism, probably the most influential version of the doctrine on record, we find that the contrast between it and panpsychism is pronounced. Nevertheless it does seem to address many of the concerns that tend to drive a philosopher to embrace panpsychism. To one who has embraced panpsychism for want of a better alternative, Russell's neutral monism may present itself as a genuine option.

Aside from the many problems involved in clearly articulating the sort of neutral monism that Russell has proposed, we find that the doctrine may face a version of the challenge that Chalmers has launched against Russellian panqualityism. Given the current state of development of Russell's neutral monism, the force of this conceivability argument is hard to assess—too many details of the view await further development.

It seems to me that the proper conclusion to draw at this point is this: The problems facing Russell's neutral monism are great but so is the promise the theory holds. The panpsychists would do well to take a serious look at that theory.¹⁵

Notes

- 1. Sometimes this view is called panexperientialism; I will follow Chalmers, and others, in taking it to be a representative of panpsychism.
- But compare William Seager's definition: "Panpsychism is the doctrine that mind is a fundamental feature of the world which exists throughout the universe" (Seager 2010). This panpsychism allows for the possibility that the fundamental is purely mental.

- 3. The resultant vs. emergent property distinction is due to Jaegwon Kim: "Resultant properties are 'additive' or 'subtractive'; for example, the weight of this table is merely the arithmetic sum of the weight of its top and the weight of its base, and hence it is predictable from the properties of its parts. For this reason, it is a resultant property" (Kim 1998, 227).
- 4. See Seager (2010) for this way of framing the dispute: given certain assumptions, the most fundamental division among theories in this area is that between panpsychism and emergentism (in both its brute and benign forms).
- 5. Chalmers uses the term 'quiddities' for the properties that play the grounding role; Alter and Nagasawa (2012) prefer the term 'inscrutables.' So, instead of 'quidditism' we might want to say 'inscrutabilism.' But 'quidditism' seems like the lesser evil.
- 6. According to panqualityism, experienced properties, not the experiences involving these qualities, function as the quiddities of basic physical entities.
- 7. Chalmers thinks that this form of neutral monism approximates the actual views of the three great neutral monists: "Something like this seems to have been the preferred view of neutral monists such as Mach, James, and Russell" (Chalmers this volume).
- 8. And if one is inclined, as Chalmers (2015) is, to identify the quiddities (in this case: simple qualities) with certain physical properties, the question whether we are dealing with a version of neutralism or physicalism becomes still more confounding.
- 9. The question whether the physical can be fundamentally mental is fraught in controversy (cf. e.g., Wilson 2006; Dorsey 2011).
- 10. Russell insists that his turn to neutral monism brings with it the rejection of sensedata: "When I was young, I thought that a sensation is essentially relational and consists of a relation between a 'subject' and a 'sense-datum.' I abandoned this view emphatically and dramatically in 1921 in *The Analysis of Mind*; and, since that time, I have not spoken of 'sense-data' unless by inadvertence. Philosophers, however, have not noticed this." (Russell 1957/1997, 305)
- 11. "We have ... a certain formal structure which is prescribed to us [by physics]. Subject to this formal structure, we are at liberty, logically, to construct our events, our points of space-time, our electrons and protons, as we please. Any set of logical constructions having the requisite formal properties is logically admissible. Mathematical logic shows an infinite number of possibilities fulfilling these conditions. The choice among possibilities, if possible at all, must come from outside; nothing in physics itself can determine our choice" (Russell 1925/1988, 277).
- 12. We have already encountered the suspicion that Russell's neutral monism is really just panpsychism under another name. The contrary view is alive and well, too. See, for example, Gregory Landini's case for the claim that "the neutral stuff are physical events in space-time; some series of which constitute material objects persisting in time and other series of which constitute minds persisting in time" (Landini 2011, 291). And by calling the events physical, Landini means that "the transient particulars are without intrinsic phenomenal character" (Landini 2011, 297).
- 13. In *The Analysis of Mind* (Russell 1921) Russell attempts the construction of numerous mental states. *The Analysis of Matter* (Russell 1927a) discusses the construction of physical entities in greater depth.
- 14. These macroquality instances—these percepts—play the role of Chalmers's macrophenomenal qualities: full-sized qualities of which the experiencer is aware.
- 15. Donovan Wishon has given me valuable comments on this paper, only a very few of which I could respond to adequately. See Wishon (2015) for a probing and nuanced account of Russell's different versions of neutral monism.

Bibliography

Alter, Torin, and Yujin Nagasawa. "What Is Russellian Monism?" Journal of Consciousness Studies 19.9–10 (2012): 67–95.

- Alter, Torin, and Yujin Nagasawa, eds. *Consciousness in the Physical World: Perspectives on Russellian Monism*. Oxford: Oxford University Press, 2015.
- Bolender, John. "An Argument for Idealism." Journal of Consciousness Studies 8.4 (2001): 37–61.
- Chalmers, David. "Panpsychism and Panprotopsychism." Consciousness in the Physical World: Perspectives on Russellian Monism. Eds. Torin Alter and Yujin Nagasawa. Oxford: Oxford University Press, 2016. 246–77.
- Chalmers, David. "The Combination Problem for Panpsychism." *Panpsychism.* Eds. Godehard Brüntrup and Ludwig Jaskolla. Oxford: Oxford University Press, this volume.
- Chalmers, David. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press, 1996.
- Clarke, David. *Panpsychism: Past and Recent Selected Readings*. Albany: State University of New York Press, 2004.
- Coleman, Sam. "Mind Under Matter." *Mind that Abides: Panpsychism in the New Millennium.* Ed. David Skrbina. Amsterdam: John Benjamins, 2009. 83–107.
- Coleman, Sam. "The Real Combination Problem: Panpsychism, Micro-Subjects, and Emergence." *Erkenntnis* 79.1 (2014): 19-44.
- Dorsey, Jonathan. "On the Supposed Limits of Physicalist Theories of Mind." *Philosophical Studies* 11.2 (2011): 207–25.
- Fodor, Jerry. Psychosemantics: The Problem of Meaning in the Philosophy of Mind. Cambridge, MA: MIT Press, 1987.
- Holt, Edwin. The Concept of Consciousness. New York: Macmillans, 1914.
- Kim, Jaegwon. Philosophy of Mind. Boulder: Westview Press, 1998.
- Landini, Gregory. Russell. New York: Routledge, 2011.
- Nagel, Thomas. "Panpsychism." Mortal Questions. Ed. Thomas Nagel. Cambridge: Cambridge University Press, 1979. 181–95.
- Russell, Bertrand. The Problems of Philosophy. London: Oxford University Press, 1912.
- Russell, Bertrand. "The Ultimate Constituents of Matter." *The Monist* 25 (1915a): 399–417. Rpt. in his *Mysticism and Logic and Other Essays.* London: George Allen and Unwin, 1917. 125–44.
- Russell, Bertrand. "Letter on Sense-Data." The Journal of Philosophy, Psychology, and Scientific Method 12 (1915b): 391–92.
- Russell, Bertrand. Mysticism and Logic and Other Essays. London: George Allen and Unwin, 1917.
- Russell, Bertrand. "Logical Atomism." First published in 1924. Rpt. in his *Logic and Knowledge*. *Essays 1901–1950*. Ed. Robert Marsh. London: George Allen and Unwin, 1956. 321–43.
- Russell, Bertrand. "The Philosophical Analysis of Matter." First published in 1925. Rpt. in his The Collected Papers of Bertrand Russell. Essays on Language, Mind and Matter 1919–26. Ed. John Slater. Vol. 9. London: Unwin Hyman, 1988. 275–84.
- Russell, Bertrand. *The Analysis of Matter*. London: Kegan Paul, 1927a. Page references to the Dover edition from 1954.
- Russell, Bertrand. An Outline of Philosophy. London: George Allen and Unwin, 1927b.
- Russell, Bertrand. "Perception." First published in 1954. Rpt. in his The Collected Papers of Bertrand Russell: Last Philosophical Testament 1943–68. Ed. John Slater. Vol. 11. London: Routledge Chapman and Hall, 1997. 304–07.
- Sayre, Kenneth. Cybernetics and the Philosophy of Mind. Atlantic Highlands: Humanities Press, 1976.
- Seager, William. "Panpsychism." Stanford Encyclopedia of Philosophy. 23 Aug. 2010. Web. 14 July 2014.
- Skrbina, David. Panpsychism in the West. Cambridge, MA: MIT Press, 2005.
- Skrbina, David. "Panpsychism." Internet Encyclopedia of Philosophy. Web. 14 July 2014.
- Skrbina, David, ed. *Mind that Abides. Panpsychism in the New Millennium*. Amsterdam: John Benjamins, 2009.

- Strawson, Galen. "Realistic Monism. Why Physicalism Entails Panpsychism." Consciousness and Its Place in Nature: Does Physicalism Entail Panpsychism? Eds. Galen Strawson et al. Exeter: Imprint Academic, 2006. 3–31.
- Stubenberg, Leopold. "Neutral Monism." *The Stanford Encyclopedia of Philosophy* (Winter 2014 edition). Ed. Edward Zalta, Web. 11 Apr. 2016.
- Stubenberg, Leopold. "Russell, Russellian Monism, and Panpsychism." Consciousness in the Physical World. Perspectives on Russellian Monism. Ed. Torin Alter and Yujin Nagasawa. Oxford: Oxford University Press, 2015.

Wilson, Jessica. "On Characterizing the Physical." Philosophical Studies 131 (2006): 61–99.

Wishon, Donovan. "Russell on Russellian Monism." Consciousness in the Physical World. Perspectives on Russellian Monism. Ed. Torin Alter and Yujin Nagasawa. Oxford: Oxford University Press, 2015. 15

Dualism and Panpsychism

CHARLES TALIAFERRO

In this essay, I propose that mental-physical, and even person-body dualism, as part of an overall theistic metaphysics, offers panpsychism a sturdy philosophical position that is more cogent than its counterparts. Both forms of dualism are supported by what I shall argue is *the primacy of the mental* over against what philosophers today demarcate as the physical. This dualist stance is able to avoid the absurdities of some contemporary physicalist accounts of mind that undermine reason itself. With theism and dualism (or theistic dualism), panpsychists can provide an account of mind and cosmos (to use terms borrowed from Thomas Nagel) where naturalists cannot; they can provide an account of the foundational trustworthiness of reason; and, paradoxically, their position is indirectly supported by the claims of some antitheistic naturalists.

At first, however, the above claims may sound naive and quite unwelcome. After all, panpsychism is itself an emerging or reemerging position that seems out of step with philosophies that give primacy to the physical sciences and wield Ockham's razor whenever possible. Why take on board dualism plus theism, two positions that receive more than their fair share of abuse in much of the current philosophical literature? My intent is to redeem dualism over against its many objections and to appeal to panpsychists' willingness to recognize the mind or the mental that goes well beyond contemporary naturalism. I provide some reasons for thinking that by taking on the ultimate mindoriented, teleological metaphysics, classical theism, panpsychism provides an eminently challenging position. (Of course, as this is an essay rather than a series of books, I will have to draw on other work to bolster some of the moves that are made in what follows.)

There are three sections that follow. In the first section reasons are given for recognizing the primacy of the mental rather than the primacy of the physical sciences. Mental causation is articulated and defended and a case is made for the nonidentity of the mental and the physical. The focus is on human persons,

but the arguments of this section will provide reasons for panpsychists to not identify the mental or psychic properties of nonhuman beings and objects. In the second section, reasons are advanced why philosophers in general, including panpsychists, should accept a dualist concept of the self as a substantial individual enduring over time, a self that is not identical with its brain or physical body. In the final section, it is argued that a theistic-dualist-panpsychist metaphysics is more cogent than its popular counterparts. Such a metaphysic is, paradoxically, supported by some of the claims of antidualist, antipanpsychic philosophers.

The two sections that follow use the term 'redeeming' with regard to dualism, as I am keenly aware that many philosophers think dualism is philosophically hopeless. My aim is to challenge this antidualist milieu, showing dualism to be plausible while at the same time showing that some popular alternatives to dualism are problematic. As the purpose of this essay is to redeem dualism for the sake of an inclusion—linking dualism, panpsychism and theism—the section headings highlight the context of the defense of dualism.

15.1 Redeeming a Dualism of the Mental and Physical for Panpsychism and Theism

Before seeking to show that panpsychists, and philosophers in general, should recognize a metaphysical distinction between the mental and the physical, I register some reservations about the term 'dualism.' None of the major figures in philosophy that are labeled as dualists used the term 'dualism.' So, you will not find the term in the work of Plato, Augustine, Descartes, Leibniz, or Locke, among others. 'Dualism' as a term was first introduced to describe Zoroastrianism with its positing of a good God and evil counterpart, and it is perhaps this association that invites some critics of dualism (especially from contemporary theologians) to think dualism is inextricably connected with treating the body as a kind of necessary evil and the soul or mind as good (see e.g., McFague 2008). Of course, the bare thesis that persons (or minds or souls or mental states) are metaphysically distinct from physical bodies carries no implications that the body is not good or that there is not a functional unity of person and body under healthy conditions (more on that below). In addition to 'dualism' being associated with a denigration of the body, what is also misleading about the term 'dualism' is that it suggests (but does not entail) that a primary conviction in a dualist credo is that there are only two kinds of things: the physical, material, corporeal, extended (etc.), and that which is nonphysical, immaterial, incorporeal, and unextended. Now, Descartes did

distinguish between the spatially extended and the nonspatially extended, but it is important to appreciate that in the Meditations, this distinction is something he arrives at by holding a positive position on the reality of himself and his thinking. I suggest that the most reasonable and charitable way to understand why someone would adopt dualism is to take seriously the primacy of our subjective, experiential awareness (and what we are aware of) and the implausibility of identifying this subjective life with what philosophers identify as the physical. In this sense, it would be more fitting to think of 'dualists' as *pluralists* as they maintain that there is more to reality than what is identified as physical. In a way, dualists may be quite agnostic about what exactly should be classified as 'physical.' All they need to do to be pluralists is to reject the forms of monism that are available or seem to be in the offing. If the term 'physical' were to evolve so as to include traditional theism, angels, souls, or Platonic entities, 'dualists' might re-describe themselves as 'physicalists' and proud monists. Short of such a wildly capacious notion of the 'physical,' 'dualism' today may be taken as the view that there are at least two kinds of things, and the mental is distinct from what most philosophers (except perhaps Galen Strawson) mean by 'physical.' Hence, in this essay 'dualism' should be read as a positive affirmation of the reality of the mental and a negative judgment that the mental is identical with the physical.

There is yet another reason to regret the use of the term 'dualism.' Many critics claim that dualism involves an extreme bifurcation of the mind and body, such that to see one's bodily behavior is not to see an embodied person but a body being moved by an incorporeal mind. Such a caricature was made famous in Gilbert Ryle's The Concept of Mind, and it has been continued by a string of philosophers including Antony Flew, Anthony Kenny, Peter Hacker, Trent Merricks, and others. What is lost is the idea that under extraordinary factors involving neurological damage or impairment of one's motor control, a person can wind up feeling like or being a kind of ghost in a machine. But, under healthy conditions, dualists can (and should) recognize that an embodied being functions as a unified being. So, if I have lost almost all motility except for the movement of my hands, to see my main body may be like seeing a container or vessel or trunk in which I have some fragmentary life; but under healthy conditions, my waving hello to you is an embodied, unified act (for a fuller development of this unity, see my 2001 article "The Virtues of Embodiment").

My case for a dualism of the mental and physical will require arguing for the epistemic primacy of the mental and for our primary awareness and exercise of what may be called mental causation. This will involve challenging the current intellectual climate.

The following description of the state of play in philosophy of mind is not quite as obvious as it was when it was written in 1998, but it still rings true in many departments of philosophy. In *Mind and a Physical World*, Jaegwon Kim writes:

The shared project of the majority of those who have worked on the mind-body problem over the past few decades has been to find a way of accommodating the mental within a principled physicalist scheme, while at the same time preserving it as something distinctive—that is, without losing what we value, or find special, in our nature as creatures with minds." (Kim 1998, 2)

This position may have to be modified somewhat, given the many arguments that have been deployed against physicalism in works such as *After Physicalism*, edited by Benedikt Paul Goecke, *Contemporary Dualism: A Defense*, edited by Andrea Lavazza and Howard Robinson, *The Waning of Materialism*, edited by Robert Koons and George Bealer, and others. But there have been, and there still are, an impressive number of philosophers who share, with Kim, a confident picture of the physical world, and a considerably less confident understanding of how to fit in what we think of as mental.

Consider two more philosophers who give primacy of intelligibility to the physical world. D. M. Armstrong offers this classic, succinct statement of his metaphysical position: "Naturalism [is] the doctrine that reality consists of nothing but a single, all-embracing spatio-temporal system" (Armstrong 1995, 35). Here is Michael Tye's position:

On the naturalist view, the world contains nothing supernatural . . . at the bottom level there are microphysical phenomena governed by the laws of microphysics, and, at higher levels, phenomena that not only participate in causal interactions describable in scientific laws but also bear the general ontic relationship to microphysical items as do the entities quantified over and referred to [in] such higher-level laws as those which obtain in, for example, geology and neurophysiology. (Tye 1994, 129)

In the wake of such positive claims about what is physical, no wonder some philosophers think that the idea of what may be nonphysical is suspect.

Kevin Corcoran and Elizabeth Anscombe are only two of the many philosophers who think that humans contain, or are, nonphysical beings. They are especially interesting as both are Christian theists, and so they cannot affirm both the coherence of theism, in which God is incorporeal or nonphysical, and the strict unintelligibility of the existence of such an incorporeal reality. Given that they already accept in their metaphysics something incorporeal, it is interesting that they find it baffling, if not incoherent, to acknowledge that there is a finite incorporeal, nonphysical reality. Here is Corcoran's position:

Try as I might I cannot bring myself to believe what my mother believes, and what many Christians down through the ages have believed, about the metaphysics of human persons. It is not that I do not understand the view. I do. So too with the traditional way of understanding the nature of human persons. I simply cannot believe that I am an immaterial thing. I can believe that some kinds of persons are immaterial – for example, nonhuman divine persons like God and the angels. But human persons like me. That I cannot believe. (Corcoran 2005, 154)

Elizabeth Anscombe suggests that it is possible that "the conception . . . of an immaterial substance . . . is a delusive one" (Anscombe 2008, 73).

But stepping back a bit, how clear a concept do we have of the physical world, and how does that match our concept of what many philosophers classify as mental: our thinking, conceiving, feeling, seeing, hearing, tasting, valuing, observing, and so on? We cannot even begin to try to understand what is physical unless we can trust our reasoning and conceptual powers, for without these we would be unable to consider whether mind-independent objects have mass, volume, size, color, odor, sound, taste, sensory qualities of heat (as opposed to heat as in mean kinetic energy), or whether the physical consists in individual things (particles) or events or fields. I suggest that the priority of intelligibility and clarity should be acknowledged as the mental, and that none of the above conceptions of the physical can be any clearer or more intelligible than the mental. This is evident when the physical is analyzed in terms of that which is intersubjective or those things which more than one person can (in principle) observe. Such an analysis must presuppose an antecedent confidence and understanding of subjectivity and observation. In fact, no statements of what is physical can be more certain than that which is mental. In developing reasons for this reversal of priorities, keep in mind that I am not at the outset presupposing or assuming that the mental is not physical. This will be argued for, but not assumed.

Let us review the earlier statements by our various physicalists or 'nearenough' physicalists. In response to Kim, I suggest that it is impossible to have a clearer conception of 'a principled physicalist scheme' than you can have of a 'scheme' which, I assume, is a concept or way of conceiving. And Kim's statement as a whole seems to commit him solidly to the reality of the mental; 'accommodating,' 'valuing,' and grasping principles are mental acts. Michael Tye writes impressively of laws of nature, and yet we can have no conception of a law of nature unless we can trust the reality and reliability of our concepts and the reality of mental causation. In this context, 'mental causation' would be evident in our grasp of laws of nature; of comprehending when it is that certain molecular, atomic, nuclear, and subnuclear events cause or explain other molecular, atomic, nuclear, and subnuclear events. We only grasp a law of nature if we can trust our reasoning, whether this is cashed out in terms of a covering law model or counterfactuals, or we adopt a philosophy of causation that recognizes basic powers. This involves the use of mental causation insofar as a person grasps the relevant causal relata, and whether the relata are immanent, located in spacetime, or transcendent and nonspatiotemporal, grasping laws of nature involves our reasoning that if certain antecedent and contemporary events obtain, then there is reason to believe this will bring about (or cause or explain) another event. The causal elements in the course of a person's reasoning may be vast and complex, but for reasoning to occur, the conclusions a person draws must (in a crucial, ineliminable way) be in virtue of their grasping the relevant premises and inferential rules. From simple mathematics to astrophysics, it is essential that we draw conclusions in virtue of grasping reasons and entailment or inferential relations.

While the following is mind-numbingly obvious, it seems to be overlooked or underappreciated: microphysics, geology, and neurophysiology cannot be practiced unless there are microphysicists, geologists, and neurophysiologists, and each of them must necessarily work with concepts, observations, and theories and be able to grasp entailment relations, the laws of logic, and so on. It is largely because of what I believe to be the ineliminability of mental causation from our reasoning that I am unimpressed by the charge that dualists face a unique, intractable problem in showing how something nonphysical can be in a causal relation with the physical. All theories of mind must find a way to secure the claim that reasoning involves mental-mental causation, that is, all theories besides eliminativism, radical behaviorism, and perhaps solipsism. In "What Is a Problem for All Is a Problem for None: Substance Dualism, Physicalism, and the Mind-Body Problem," Kenneth Himma points out that philosophers who embrace the causal closure of the physical (and thus only allow for physical-physical interaction) may have avoided one problem, but then they have the problem of justifying why they rule out in principle the causal role for the hurtfulness of pain.

Eliminating the ontological distance between mental states and brain states by reducing the former to the latter solves the conceptual puzzle of how the mental and physical can interact because there is no conceptual mystery about how one physical state can cause another. But it solves this puzzle only by ruling out, as a conceptual or nomological matter (depending on the character of the reduction), any causal role for the hurtfulness of pain. (Himma 2005, 88)

Unless one takes the radical move of evading the evident causal role of the mental (something I believe Dennett does), all theories of mind must find a way to secure mental causation as a reality or they risk undermining the normativity of reason itself.

I suggest that the proper way to approach mental-physical causation is not to assume at the outset that we have a foundational, primary understanding of physical causation, but to assume, rather, that we grasp mental causation. We should then be thinking of the problem at hand in a way that turns Kim's position around: we know that we have exercised our thinking, observing, and experimenting to arrive at a valuable scientific understanding of the world. How then should we accommodate this view of the world into a comprehensive philosophy of science and nature? This question may require both science and philosophy, but in so doing, the last resort should be doubting that we can reason, and thus doubting the existence of what I am calling mental causation. This would put us in the unenviable position of providing reasons as to why we do not reach conclusions in virtue of grasping reasons. In light of this, it would seem to be spurious or at least unwise to assume at the outset that we know the necessary conditions for causal relations in a way that would rule out mental causation for any of the major extant theories of mind. Besides, some of the objections to dualist mental-physical interaction rest on conceiving of the mental as nonspatial, whereas various dualists historically and in modern times have thought of the mental as spatial (e.g., Henry More, Henry Price, and George Moore).

Continuing the engagement with our initial cast of naturalist-materialists, how certain can one be that "reality consists of nothing but a single, allembracing spatio-temporal system" (Armstrong 1995, 35)? I am not sure, but I suggest that it is more certain that reality consists of persons who can propose extraordinary theories about spatiotemporal systems. If we know anything about reality at all, I suggest the latter is our surer thesis rather than Armstrong's proposal. Also, as noted above, significant dualists have affirmed the spatial extension of the mental (sensory, visual fields, one's felt body or body image in which one feels oneself as spread out in space) and it is not clear whether materialists want to think of sounds, smells, thoughts, propositions, properties, and mathematical entities as spatial.

Returning to Corcoran and Anscombe, what would it be to think of oneself as an immaterial thing or to contend that an immaterial substance is a delusive idea? I suggest taking one step back to consider a different question. Whether you or anything is nonphysical or immaterial, isn't it evident (or isn't there good reason to believe) that persons are substantial individuals who endure over time? I assume that Corcoran believes he and his mother retain personal identity over time. And I assume that Anscombe is certain that when she thinks about immaterial substances, it is her, Anscombe, who is making a proposal about metaphysics and intelligibility to others. Let's come back to self-awareness, instead of the case of Corcoran and Anscombe that we will address in the next section, and reemphasize briefly the order of reasoning in the case for dualism: Dualists do not (typically) begin by proposing the hypothesis that persons are immaterial or incorporeal. They propose that persons exist or that there are conscious, experiential states, and then look at the physicalist proposals (either in terms of identity, reductive or nonreductive terms) and find them wanting.

So, given the thesis shared by dualists and panpyschists, but not by eliminativists, that the mental exists, how plausible is it that the mental is strictly identical with the physical? If the physical is what philosophers like Daniel Dennett, Michael Tye, the Churchlands, and others of a like mind think it is, then I believe the knowledge argument gives us abundant reason not to strictly identify the mental with the physical. The knowledge argument, as developed by many philosophers such as Frank Jackson, Thomas Nagel, Richard Swinburne, and Timothy Sprigge, has different versions but basically presses home the point that, if the mental and physical are identical, then to know of one is to know of the other. I believe the argument is best formulated in terms of particular or token identity statements in a way that makes clear the evident phenomenological grasp that we have of what it is we think, experience, feel, decide, and so on. It is the claim about having a phenomenological grasp of the mental that is key. Alvin Plantinga, for example, offers the following reason for not identifying the mental and physical. I find it compelling but in need of a little phenomenological extension.

Presumably neither [electrons nor quarks] can think – neither can believe, doubt, want, fear, or feel pain. But then a proton composed of quarks won't be able to think either, at least by way of physical relations between its component quarks, and the same will go for an atom composed of molecules, and an organ (e.g. a brain) composed of cells. If electrons and quarks can't think, we won't find anything composed of them that can think by way of the physical interaction of its parts. (Plantinga 2008, 53)

Without appealing to phenomenology, this reasoning may seem like the fallacy of composition, according to which a whole may have properties and functions

not possessed by any of its parts. To take a mundane example, the parts of a bicycle are not bicycles, but put them together in the right way and presto, you've got a bike. But Plantinga's reasoning is sound once we make clear that feeling pain, reasoning, doubting, wanting, and fearing are all experiential states and activities that we can grasp and identify, knowing what they are like. And one can both know all about one's mental life without knowing anything of the physical constitution of the brain, one's anatomy, and so on, and know all about the physical processes in the brain without knowing anything of the content of a person's mental life.

There are, of course, hundreds of objections to the above argument, and hundreds of replies. But so far, I think none of the rejoinders by physicalists are successful. One popular way to attempt to defuse the argument is to cast it as making only an epistemological point with no metaphysical consequences. But here, putting the point somewhat awkwardly, what we know of thinking is not something that can be negated by its object turning out to be something other than thought. Thinking is thinking; it is real and not merely a reflection of the way we employ mental concepts to get at what is actually real. Using different terms, the feeling of pain, being afraid, doubting, and so on, are ways of experiencing and responding to various things, and their essence lies in the experiencing itself. Take away experience, and you take away feeling pain et al. Consider three objections before considering why panpsychists can use some help from dualism.

Objection 1. A radical challenge would be to deny that our grasp of the mental is at all reliable. Might it be the case that we only have the illusion of being conscious? Nicholas Humphrey (Psychology, London School of Economics) seems to adopt this position. I cite him at length rather than try to paraphrase his stance:

I believe that human consciousness is a conjuring trick, designed to fool us into thinking we are in the presence of an inexplicable mystery. Who is the conjuror and why is s/he doing it? The conjuror is natural selection, and the purpose has been to bolster human selfconfidence and self-importance—so as to increase the value we each place on our own and others' lives. If this is right, it provides a simple explanation for why we, as scientists or laymen, find the 'hard problem' of consciousness just so hard. Nature has meant it to be hard. . . . Can I prove it? It's difficult to prove any adaptationist account of why humans experience things the way they do. But here there is an added catch. The Catch-22 is that, just to the extent that Nature has succeeded in putting consciousness beyond the reach of rational explanation, she must have undermined the very possibility of *showing* that this is what she's done. But nothing's perfect. There may be a loophole. While it may seem—and even be—impossible for us to explain how a brain process could have the quality of consciousness, it may not be at all impossible to explain how a brain process could (be designed to) *give rise to the impression* of having this quality. (Consider: we could never explain why 2 + 2 = 5, but we might relatively easily be able to explain why someone should be under the illusion that 2 + 2 = 5). (Humphrey 2005)

Reply: It seems that Humphrey is claiming that we may have impressions, beliefs, thoughts, or experiences of consciousness but are fooled by these into thinking that we are conscious. But if we have impressions, beliefs, experiences, and thoughts (however they are brought about), then, ipso facto, we have conscious states and are conscious. And at the risk of seeming thoroughly Cartesian, how can a person be fooled unless they are led to believe something false or misleading? For consciousness, or the hypothesis that the brain produces consciousness to be akin to 2 + 2 = 5, then the idea that consciousness exists and the hypothesis that it emerges from brain processes would be necessarily false. But how can we think anything is necessarily false or impossible or necessary . . . unless we can think and are conscious? The thesis that brain processes could "have the quality of consciousness" (Humphrey 2005) may make sense, but it appears to be a category mistake and at odds with what we know in self-awareness (more on that in the next section). It seems that there cannot simply be consciousness—there must be a subject or individual that is conscious. And processes are not individuals.

The idea that natural processes may have produced us with completely unreliable cognitive access to reality, including the reality of our own states of mind, is an interesting thesis, though it would be hard to argue for or to know about this unreliability unless we had some minimal cognitive competence.

Objection 2. The claim that we know our mental life, our thinking, feeling, schemes, and so on, better than we know what philosophers take to be physical is wide open to challenge. We may know what and why we are thinking along certain lines, but we may have no idea what thoughts are. What are they made of? Where did they come from? Where do they go when we are not thinking about them? Don't we know more about water and its constitution than we do about almost any mental state?

Reply: We can and do know a great deal about the mental and, indeed, we know more about the mental than we know about water. To know anything about water, you need to have abundant concepts and thoughts about water's constitution, its distribution, its properties at different altitudes, and so on. Asking what thoughts are made of or their origin and destiny are fine, but all

this involves, and must involve, confidence that we can indeed think and that we know how to think. We do not know what stuff thoughts are made of in the sense of knowing what clouds are made of, but we can and obviously do know indefinitely many things about thoughts and thinking. You cannot think of three things, until you think of two. And to claim that one has no idea what thinking is is itself a line of thinking, and a very peculiar one. It is peculiar because if you do not know what thinking is, how can you be so sure of what you are thinking or not thinking? I suggest that one cannot know more about water than about the mental because all your thought about water involves an antecedent, prior mental awareness, and to forget that it is your mental awareness that allows you to know anything about water would be like forgetting that you need to move if you are going to go for a run.

Objection 3. While a natural second objection might appeal to the private language argument inspired by Wittgenstein, I have addressed the argument elsewhere (see Taliaferro 1994; Goetz and Taliaferro 2011) and so will use instead the following objection to advance something very close to, but not identical with, the private language argument. Some philosophers have argued that we have a clearer grasp of mind-independent physical objects than we do of our mental states such as the appearance of physical objects. Rob Lovering offers a recent, vivid version of this position:

I find it nearly impossible to believe that being guided by beliefs regarding mere appearances rather than being guided by beliefs, regarding the real existence of physical objects—including not only with ducks, but stampeding horses, swinging swords, incoming spears, falling rocks, food, doors, etc.—is conducive for successfully navigating through life. After all, mere appearances are just that, mere appearances. So, unless one also thinks that appearances give us good reason to believe that physical objects really exist, I fail to see how mere appearances could serve one in successfully navigating through life. "I am being appeared to sword-swingingly; thus I ought to move out of the way" makes sense if the appearance of a swinging sword gives one good reason to believe that the swinging sword really exists. (Lovering 2013, 82)

Reply: I do not question whether the appearance of a sword is (other things being equal, e.g., one is not in the mental state of Lady Macbeth) a good reason to believe that there is a sword, and that "navigating through life" (Lovering 2013, 82) under healthy conditions, it would be unwise to routinely examine one's appearances and raise philosophical questions about their reliability. But it seems unintelligible to think you might be in a sword

fight and avoid being stabbed without making excellent use of how your opponent appears to you. Perhaps the problem in Lovering's examples is the use of the term mere; he seems to suppose that if we say X appears or X appears to be the case, this implies that this is a mere, or only an appearance. This recalls older debates about whether knowledge involves belief on the grounds that if you claim to believe X, persons sometimes assume you only believe but do not know that X. As in the case of knowledge, 'belief' and 'appearance' need not be given a minimal interpretation. The notion of reliable and accurate appearances—so accurate and reliable that a prudent person would be better off not questioning them-makes perfect sense, as does the idea that you know something when you believe it plus certain other conditions are satisfied (e.g., your conclusions do not involve essential reasoning relying on a false premise). Lovering's position seems to rely on misleading examples. A misleading example would be to infer from someone saying 'Eric and Miriam do not live in a house. They live in a mansion!,' that a mansion is not a house. Mansions are houses; they are simply not 'mere' houses, but houses of grandeur (see Lehrer 1990, 27).

15.2 Redeeming a Dualist Conception of the Self for Panpsychism and Theism

So far, the first section may be supportive of 'merely' property dualism or maybe even idealism, as I have done little to bolster the idea that we either have a noncontroversial understanding of the physical or that there are good reasons one should countenance a substantial self that is not identical with physical things and processes.

Following the lead of the first section, let us first consider the extent to which our first-person self-awareness gives us good reason to believe that we are substantial selves who endure over time. Assuming we are successful in finding such good reason, let us then see whether or not it is reasonable to identify ourselves with our bodies or a part of our body (e.g., brain). I shall argue that such an identity is not plausible, and thus panpsychism is best seen in partnership with a dualist account of the self.

In keeping with recognizing the primacy of the mental and mental causation in the first section of this essay, I propose that it is evident or more evident that you and I are selves or persons, beings who endure over time and who think, feel, and act, than it is that you and I are not selves or persons and do not endure over time, think, feel, or act. As Charles Campbell once remarked, to hear Big Ben ring four times, you have to have heard it ring three times. I take it that philosophical treatments of persons that deny this are implausible or must be supported by a powerful argument. Efforts to dissuade us from this evident first-person experience can bring out the implausibility of such a denial.

Consider briefly three efforts to dislodge the evident awareness we have of ourselves as selves or persons. Peter Geach has claimed that the use of "I" in "soliloquies . . is redundant and has no special reference; 'I am very puzzled at this problem' really says no more than 'This problem is puzzling'" (Geach 1957, 120). Moritz Schlick proposed, "Descartes had no right to start his philosophy with 'I think,' instead of saying 'it thinks'" (Schlick 1949, 166). Third, consider Bertrand Russell's proposal that the "I think" of Descartes can be paraphrased as "there is thinking" (Russell 1945, 567).

I suggest that none of these efforts to evade the first-person point of view is successful. Contra Geach, when I say 'this problem is puzzling,' I may not personally be at all puzzled. I may simply be observing that others find the problem baffling or paradoxical. It would also be hard to imagine that anything could be puzzling if there were no (at least in principle) people who are or would be puzzled. Schlick's proposal seems pale and disturbing. What would it be like for me to assert that 'It thinks'? Off hand, my remark would conventionally be interpreted as my referring to some other thing (e.g., a nonhuman animal as in 'it thinks I will give it a treat'). If the remark is otherwise meant to capture (or eliminate) the first-person point of view, it would seem to fail, for a speaker's use of indexicals like 'it,' 'here,' 'there,' and so on, is best interpreted as designating something the speaker (the subject) is drawing our attention to. And in the case of 'it thinks,' it seems more likely to interpret the speaker as referring to something inside her or a part of her (the brain? a fetus?), rather than referring to herself. Russell's proposal seems to land us in the absurd position of thinking without a thinker, which seems no more intelligible than swimming without a swimmer or seeing without a seer.

In *Naturalism and the First-Person Perspective*, Lynne Baker forcefully brings to light the formidable difference between knowing that 'Lynne Baker will die one day' and expressing her first-person awareness when she says 'I, Lynne Baker, myself, know that I will die one day.' Baker argues that even if one were to eliminate reference to 'I' in Descartes' reasoning, this would not show that first-person awareness can be eliminated in such sentences as "I regret that I find this puzzling" (Baker 2013, 38).

If the first-person point of view and our awareness of ourselves as subjects are so stubborn, why are some philosophers so highly motivated to tame or evade them? I suggest it is partly because it appears that if you accept the evident self-awareness of subjects, you may be tagged with some form of dualism. The knowledge argument can be brought to bear on matters: If the subject or person is their brain or their physical body as a whole, then to observe one is to observe the other, but no amount of study of the brain and body as

'unproblematic physical objects' (to borrow a phrase from Paul Churchland) will disclose anything about a person's mental life or self-awareness. As Dennett puts the point neatly, "The trouble with brains, it seems, is that when you look in them, you discover that there's nobody home" (Dennett 1992a, 29). And Dennett states in another essay: "'Where are these selves?' he asks rhetorically. It is a category mistake to start looking around for them in the brain" (Dennett 1992b, 109). This is not a problem that can be overcome by invoking different points of view, as in cases of when the same thing may be known in two different ways. You might know of Mohammad Ali and not know that he is also called (or was called) Cassius Clay, or you may know about water without realizing it is H₂O. In the case of the boxer, we have two distinct but complementary perspectives, and in the case of water, there is a difference between identifying a liquid and identifying its composition; no philosophical puzzles arise. Once this is clarified, one may readily grasp that to see Mohammad Ali is to see Cassius Clay, and to see water is to see H₂O. But there is no such confluence in the case of the mental and physical. We only know through neuroscience which conscious states correlate with which brain states by matching our observations of physical states with what subjects report about their experiences.

If we seem to have good reason to believe that we are subjects, and that we are not identifiable with our brains or physical bodies, then it seems reasonable to think that we as subjects are not identical with our brains or bodies, though under healthy conditions the person and body function as a unity. Back to Corcoran and Anscombe as discussed in the first section: dualists do not require a clear account of what it is to be immaterial or nonphysical. All that is needed is the negative judgment that persons are not metaphysically identical with their physical bodies. If pressed, however, one can offer a highly detailed account of a world without physical objects: this can be provided by an idealist account of our present world in which there are no mind-independent or non-mind-constituted objects (see Taliaferro 1994).

Dennett is quite clear in terms of his priorities. The only explanations he is prepared to accept are those that explain away the subject:

If we are to explain the conscious Subject, one way or another the transition from clueless cells to knowing organizations of cells must be made without any magic ingredients. This requirement presents theorists with what some see as a nasty dilemma. . .. If you propose a theory of the knowing Subject that describes whatever it describes as like the working of a vacant automated factory—not a Subject in sight—you will seem to many observers to have changed the subject or missed the point. On the other hand, if your theory still has tasks

for a Subject to perform, still has a need for a Subject as witness, then ... you have actually postponed the task of explaining what needs explaining. To me, one of the most fascinating bifurcations in the intellectual world today is between those to whom it is obvious—obvious—that a theory that leaves out the Subject is thereby disqualified as a theory of consciousness (in Chalmers' terms, it evades the Hard Problem), and those to whom it is just as obvious that any theory that doesn't leave out the Subject is disqualified. I submit that the former have to be wrong. (Dennett 2005, 145)

There is not space to explicate and critique Dennett's idea that the self is a narrative construct rather than an irreducible self who thinks, talks, writes books, and so on. It will have to suffice here to claim that Dennett certainly seems to explain the self or subject in terms that resemble a vacant factory, not a subject in sight, and that there is a heavy burden of proof on those who deny the reality of the self as a substantial individual enduring over time who thinks, and so on. In part, I think Dennett's position is more illuminating of the self and consciousness than almost any other philosopher's because he seems to take seriously the cost of rejecting dualism. In a sense, Dennett's work makes the case that if you recognize the self as a substantial individual with enduring powers of thought, action, and so on, you are driven to dualism. If you wish to avoid dualism, you should deny the self as an enduring, real individual.

So, if a dualism of person and body is reasonable or acceptable, or at least more reasonable than its critics acknowledge, then I suggest that panpsychists adopt a duality of the mental and physical as well as the metaphysical distinction between person and body. Let us now consider a combined panpsychismdualism stance.

15.3 Panpsychism and Dualism Together with a Theistic Metaphysics of Nature

As noted at the beginning, the first two sections contained the word 'redemption' in the title since some work had to be done to fend off the immense underestimation of the plausibility and promise of dualism in order for pansychists to welcome dualists and accept a dualist account of psychic or mental properties. I have not defended panpsychism here, for that is in the very capable hands of others in this book. I suggest that dualism (D) and panpsychism (P) have much more in common than the physicalists we have considered so far, from Kim to Dennett. Both D and P recognize the reality of the mental (or related terms like psychic or psychological or experiential properties). Advocates of D and P resist reductive accounts that wind up, in the end, as eliminative, and they tend to be unimpressed by accounts that treat consciousness or the mental as an irreducible reality that is emergent from nonconscious mental elements. Interestingly, Galen Strawson uses the same term to disparage emergentist accounts as Dennett uses to castigate accounts that leave us with an irreducible self; both philosophers accuse their opponents of appealing to magic (see Strawson 2006; Dennett 2005). It is also interesting that Dennett regularly links his opposition to dualism with his opposition to theism (in most of his books after *Consciousness Explained*). This is a good lead-in for considering a comprehensive foundation for P and D. That foundation is classical theism.

In what follows I will be relying on work I have done elsewhere, as there is simply no way to articulate and defend the claims I will be making about theism in what follows (see e.g., Taliaferro 1994; Taliaferro and Evans 2011). But I believe enough will be clear for assessing the overall philosophical strategy.

Classical theism is the ultimate reversal of Dennett's strategy of explaining mind in nonmental, ultimately nonconscious terms. In classical theism, the existence and continuation of the cosmos is sustained in being by God, the necessarily (noncontingent) reality of power, goodness, and knowledge, a being that is eternal (or everlasting) and omnipresent. In theism, there is an account of why the cosmos exists and continues in existence, why there is a cosmos sufficiently stable for there to be stars, planets and, at least on our planet, the emergence of life, consciousness, values, scientists who can apply their inquiry into the structure of reality, and so on (see Collins 2006 for a detailed exposition). Putting the theistic point of view in Dennett's terms, such an account basically explains creatures with minds in terms of a transcendent, divine mind. There is no need for theists to explain the emergence of mind from nonmind, as the mind of God is the necessarily existing foundation of the existence of all contingent things. Classical theism provides a teleological explanation of why our cosmos exists whereas naturalism (in virtually all its forms) does not offer a reason for why there is a cosmos at all or why our structured, turbulent (but stable enough for galaxies) cosmos exists and continues to exist (see Goetz and Taliaferro 2008; Taliaferro and Evans 2011). Theism therefore can provide an account of why there is a panpsychic cosmos and why there should be creatures with minds and different levels of consciousness pervading the cosmos.

A dualist-panpsychic theism is, I suggest, paradoxically supported by some of the claims of naturalists like Dennett. As we have seen, Dennett privileges the physical sciences, and proposes that they provide the model of explanation for all things. Here is a much-cited proposal by Dennett: The prevailing wisdom, variously expressed and argued for, is materialism: there is only one sort of stuff, namely *matter*—the physical stuff of physics, chemistry, and physiology—and the mind is nothing but a physical phenomenon. In short, the mind is the brain. According to the materialists, we can (in principle!) account for every mental phenomenon using the same physical principles, laws, and raw materials that suffice to explain radioactivity, continental drift, photosynthesis, reproduction, nutrition, and growth. (Dennett 1992, 33)

Even if (in principle) the mind is the brain, we cannot begin our philosophical inquiry with the brain. As argued earlier, if we searched the brain without relying on the testimony of subjects as to their experiential states, we would have no knowledge of the mental states correlated with neural states. Moreover, the practices of physics, chemistry, and physiology, as well as the explaining of radioactivity, continental drift, photosynthesis, reproduction, nutrition, and growth would not be intelligible or possible unless we accepted mental causation. Physics, et al, involve our grasping reasons on the basis of observation and reflection, along with entailment relations, that allow us to conclude (to use one of the most simple inferences possible) that if A causes B and A occurs, then, ceterus paribus, A causes or will cause B to occur. The events and laws Dennett valorizes do not involve any essential reference to the mental. If scientists are only to appeal to the kinds of laws that account for continental drift in their account of human reasoning and consciousness, then it seems that human reasoning and consciousness are not in the offing, except as initial descriptions of events that will, in the end, be cast off or explained away.

We can all share with Dennett a valorization of the physical sciences with the proviso that the practice of physical science itself presupposes mental causation, which would lead us to a nonreductive and ultimately dualistic understanding of the mental and of persons.

The aim of this essay is ambitious and other work will need to be drawn on to fill out the case before you (see Taliaferro and Evans 2011). But I hope enough has been sketched to see the appeal of a combined panpsychic-dualisttheist position. This provides an account of the cosmos whereas naturalism does not; it is able to provide an account of the trustworthiness of our mental causation which is essential in any reasoning whatever; and it avoids the absurdities or counterintuitive results of seeking to eliminate the self, the mental, and reasoning that philosophers like Dennett feel compelled to embrace in order to avoid dualism and theism.

Bibliography

- Anscombe, Elizabeth. "The Immortality of the Soul." *Faith in a Hard Ground*. Ed. Mary Geach and Luke Gormally. Exeter: Academic, 2008. 69–83.
- Armstrong, David. "Naturalism, Materialism, and First Philosophy." Contemporary Materialism. A Reader. Ed. Paul Moser and J. D. Trout. London: Routledge, 1995, 35–50.
- Baker, Lynne. Naturalism and the First-Peron Perspective. Oxford: Oxford University Press, 2013.
- Collins, Francis. The Language of God. New York: The Free Press, 2006.
- Corcoran, Kevin. "The Constitutional View of Persons." *In Search of the Soul*. Ed. Joel Green and Stuart Palmer. Downers Grove, IL: Intervarsity, 2005. 153–88.
- Dennett, Daniel. Consciousness Explained. Boston: Back Bay Books, 1992a.
- Dennett, Daniel. "The Self as a Center of Narrative Gravity." Self and Consciousness: Multiple Perspectives. Ed. Frank Kessel, Pamela Cole, and Dale Johnson. Hillsdale, NJ: Erlbaum, 1992b. 103–15.
- Dennett, Daniel. Sweet Dreams: Philosophical Obstacles to a Science of Consciousness. Cambridge, MA: MIT, 2005.
- Geach, Peter. Mental Acts. London: Routledge, 1957.
- Goecke, Benedikt, ed. After Physicalism. Notre Dame: University of Notre Dame Press, 2012.
- Goetz, Stewart, and Charles Taliaferro. Naturalism. Grand Rapids: Eerdmans, 2008.
- Goetz, Stewart, and Charles Taliaferro. A Brief History of the Soul. West Sussex: Blackwell, 2011.
- Himma, Kenneth. "What Is a Problem for All Is a Problem for None: Substance Dualism, Physicalism, and the Mind-Body Problem." American Philosophical Quarterly 42.2 (2005): 81–92.
- Humphrey, Nicholas. "What Do You Believe Is True Even Though You Cannot Prove It?" *The* 2005 Edge Question 2005: n. p. Web. 25 Sept. 2013.
- Kim, Jaegwon. Mind in a Physical World. Cambridge: MIT Press, 1998.
- Koons, Robert, and George Bealer, eds. The Waning of Materialism. London: Routledge, 2010.
- Lavazza, Andrea, and Howard Robinson, eds. *Contemporary Dualism: A Defense*. London: Routledge, 2014.
- Lehrer, Keith. Theory of Knowledge. London: Routledge, 1990.
- Lovering, Rob. God and Evidence: Problems for Theistic Philosophers. London and New York: Bloomsbury Academic, 2013.
- McFague, Sallie. A New Climate for Theology: God, the World, and Global Warming. Minneapolis: Fortress Press, 2008.
- Plantinga, Alvin, and Michael Tooley. Knowledge of God. Oxford: Blackwell, 2008.
- Russell, Bertrand. A History of Western Philosophy. New York: Simon and Schuster, 1945.
- Schlick, Moritz. "Meaning and Verification." *Readings in Philosophical Analysis*. Ed. Herbert Feigl and Wilfrid Sellars. New York: Appleton-Century-Crofts, 1949. 146–70.
- Strawson, Galen. "Realistic Monism: Why Physicalism Entails Panpsychism." Journal of Consciousness Studies 1310–11 (2006): 3–31.
- Taliaferro, Charles. Consciousness and the Mind of God. Cambridge: Cambridge University Press, 1994.
- Taliaferro, Charles. "The Virtues of Embodiment." Philosophy 76.1 (2001): 111-25.
- Taliaferro, Charles, and Jil Evans. *The Image in Mind: Theism, Naturalism, and the Imagination*. London and New York: Continuum, 2011.
- Tye, Michael. "Naturalism and the Problem of Intentionality." *Midwest Studies in Philosophy* 19.1 (1994): 122-42.

16

Idealism and Panpsychism

UWE MEIXNER

16.1 Varieties of Panpsychism

Consciousness is usually considered to be something so new, so different from what it came from, that its emergence appears to be stunning if not miraculous. Philosophers are rather uncomfortable with *miracles*—and *wonder* is felt to be more of a burden than a joy. Thus, the proposition 'first there was nothing *psychical* in the world, then there was consciousness' has seemed to many philosophers an assertion that is hard to swallow. Panpsychism comes to the rescue. For panpsychism is the doctrine that the psychical is ubiquitous in the world, whether the world is considered synchronically or diachronically, macroscopically or microscopically. According to panpsychism, there is a psychical *aspect to everything*, whether past, present, or future; whether at the micro- or the macrolevel of existence. The great general advantage of this doctrine, if believed, is that it mitigates, right away, the discomfiting philosophical wonder one feels at the emergence of 'concentrated' consciousness. And perhapsone is inclined to hope—if developed and integrated with other theories, it would even be able to dispel that wonder entirely. This is the promise of panpsychism. The great general *disadvantage* of panpsychism is that, unfortunately, it appears to be far less credible than what it is supposed to help make more rationally comprehensible (if not acceptable): the coming into being of 'concentrated' consciousness.

Panpsychism is incompatible with physicalism. This does not mean, however, that it is a form of dualism. There are both dualistic and idealistic forms of panpsychism, just as there are both holistic and atomistic forms of it. Spinoza's metaphysics (as presented in the *Ethics*) offers an example of *holistic dualistic panpsychism*: the psychical dimension (or 'attribute') of God, or Nature, matches God's physical dimension in all of its parts, is parallel to it throughout.¹ Hume's metaphysics (as presented in the *Treatise*) offers an example of
atomistic idealistic panpsychism: the impressions and ideas that make up everything (from material objects to selves) are purely psychical elements. In contrast to these historical examples, modern panpsychists opt for *atomistic dualistic panpsychism* (though they usually do not use the evil terms 'dualistic' or 'dualism,' and even try to dress it as a form of *monism*). This view says that there is a psychical (or 'proto-psychical') aspect to every elementary particle, and that macroscopic psychical phenomena—such as 'concentrated' consciousness—are attached to (i.e., are a determined function of) those huge, organized aggregations of elementary particles which are the bodies of living animals.²

There is a form of panpsychism which is seldom (if ever) explored: *holistic idealistic panpsychism*. In my view, holistic idealistic panpsychism is the best option for a philosopher determined to be a panpsychist. This is so because it is able to avoid defects that the three other forms of panpsychism cannot get rid of.

16.2 The Defect of Dualistic Forms of Panpsychism

The defect of dualistic forms of panpsychism is simply this: they do not really solve—or even help to solve—the problem, which they are intended to solve. They do not make dualism more palatable, let alone just as pleasing to the metaphysical mind as monism. Concerning the *holistic* option for dualistic panpsychism: if the dimension of the physical is in its entirety paralleled by the dimension of the psychical (as in Spinoza's philosophy), then this merely makes a riddle universal—pan-enigmatic, so to speak which beforehand was only local: How is it, how can it be that the psychical matches the physical? And concerning the atomistic option for dualistic panpsychism: from a rational point of view, it is rather curious that one finds the presence of a psychical aspect in an elementary particle less surprising than the presence of that aspect in a living animal body with a fully developed nervous system.³ The best option for dualists is, therefore, not to become panpsychists, and to remain emergentists. Presupposing that dualism is presented as a naturalistic option (which it needn't be, but can be), psycho-physical emergentism involves the belief in psycho-physical, or physico-psychical, laws of nature which, if the right circumstances come about, will produce the psychical phenomena (and mainly, of course, consciousness) automatically, 'inexorably'-not from nothing, but out of the potentiality of matter (a reservoir which holds, as we know, countless utterly surprising things).

It is perhaps not amiss to remark, in this context, that the potentiality of matter is not an unseen tiny actuality in microscopic—subatomic—matter. For example, it is within the potentiality of matter that iron rusts and gives off heat. But neither the iron atoms nor the elementary particles from which they are built rust or give off heat. There simply is no tiny, microscopic actuality from which the big, macroscopic actuality is constituted by appropriate aggregation (in fact, in this case, there *cannot* be such a microscopic actuality). Nevertheless, iron rusts and gives off heat-because it can do so (that is, it is within the potentiality of matter) and because the laws of nature *require* it to do so if the right circumstances come about. By the same token, organic tissue brings forth consciousness—because it can do so (it is within the potentiality of matter) and because the laws of nature *require* it to do so if the right circumstances (including inner, organizational circumstances) come about. The fact that I am conscious entitles me no more to believe that the atoms of my body or of my nervous system have a tiny bit of consciousness (or 'protoconsciousness') attached to them than the fact that this piece of iron gives off heat entitles me to believe that its atoms have a tiny bit of heat attached to them.

It is evident that the general assumption is quite unwarranted, indeed false, that, for any predicate F, the being-F of macroscopic actualities that are F is best explained by the organized aggregation of microscopic actualities that are F. Some particular instances of the assumption (regarding some specific predicate F or other) are not obviously false but still unwarranted and positively bizarre if it turns out that those who adopt those instances (into their belief-system) do not even know whether there are microscopic actualities with the (relevant) predicate F and what being F would even mean for microscopic actualities. An instance of the above general assumption is, in fact, adopted by atomistic dualistic panpsychists (the mainstream panpsychists): just substitute 'conscious'-or, alternatively, 'with a psychical aspect'-for 'F'. Atomistic dualistic panpsychists freely admit that they do not even know whether there are microscopic actualities with a psychical aspect; that they do not even know what being with a psychical aspect (let alone, *being conscious*) would mean for microscopic actualities. However, they are not at all bothered by these, as one would think, embarrassing admissions. Their nonchalance may seem surprising, but in fact it is not. For that there are and indeed must be microscopic actualities with consciousness or a psychical aspect is, in the end, a 'fact' that atomistic dualistic panpsychists infer-via an inference to the best explanation. They base this inference on the very assumption (which they take to be an a priori *truth*) that the fact that macroscopic actualities have consciousness or a psychical aspect is best explained by the organized aggregation of microscopic actualities with consciousness or a psychical aspect.

16.3 Idealism—the Other Monism

For those who are uncomfortable with straight dualism, there is a better form of panpsychism than dualistic panpsychism: one can be a panpsychist and abandon dualism altogether. This other view is *panpsychistic idealism*⁴—which designation is, in fact, a tautology, just like 'unmarried bachelor' or 'female mare.' For how could *idealism*—it is *ontological* idealism we are talking about—*not* be panpsychistic?⁵ It is true that idealism is unfashionable these days (a few hundred years ago it ruled the roost, just like physicalism does today); but that should not detain a panpsychist from adopting it—provided, of course, that the arguments against idealism prove insufficient.

In fact, it seems that there are more prejudices against idealism than arguments (thus idealism suffers the same fate as dualism). Here are some of those prejudices: (1) Idealism denies the existence of the physical. (2) Idealism proposes that reality depends in all its aspects on the human will. (3) Idealism entails solipsism. (4) Idealism contradicts the testimony of our senses (are there not things I can bump into?). (5) Idealism is incompatible with science and, to boot, religion. None of these prejudices withstands scrutiny. A mature form of idealism—which can be found, for example, in the works of Edmund Husserl—is compatible with science and religion, does not contradict the testimony of our senses, does not entail solipsism, does not propose that reality depends in all its aspects on the human will, and does not deny the existence of the physical.

The basic onto-epistemological fact that underlies idealism is the fact that

(I) *the world for us* is in its entirety an object of our consciousness.

Something that does not enter in any way into our consciousness remains *nothing* for us. Adapting one of Wittgenstein's apothegms, one might also say, the limits of our consciousness mean the limits of our world.⁶ This is fairly trivial. An entirely nontrivial ontological thesis of idealism results if one drops in the phrase 'the world for us' the words 'for us,' obtaining from (I):

(II) *The world* is in its entirety an object of our consciousness.

This transition from the fairly trivial onto-epistemological thesis to the entirely nontrivial *ontological* thesis is, of course, not a logical inference (since it is not a logical truth that *the world* is identical to *the world for us*); and yet it is not a fallacy, either. It is not a fallacy because one can very well argue that whatever difference there may be between the world and the world for us, that difference is just nothing to us. Hence the difference should not be assumed to exist and might as well be assumed not to exist—in application of Ockham's Razor in a generalized form (in which that principle also demands that explanations and differentiations should not be multiplied unnecessarily). A skeptic may well point out that *termites*, if they had the intelligence, could argue on this very basis that *the world* is identical to *the world for termites*—and would be utterly wrong (though they would never notice it). But this is a mere dramatization of the previous observation that it is *not* a logical truth that the world is identical to the world *for us*. Although there is no logical or rational compulsion to assume the identity in question, it might, for all we know, be *true*.

Nevertheless, the credibility of idealism can be greatly increased if one makes the transition from us to all conscious beings. Then the basic onto-epistemological fact is this: the world for the conscious beings is in its entirety an object of the consciousness of the conscious beings. And the ontological thesis of ontological idealism most likely to be true results if one replaces 'the world for the conscious beings' by 'the world': the world is in its entirety an object of the consciousness of the conscious beings. This ontological thesis does not follow logically from the onto-epistemological thesis, but the transition can be justified on the basis of the *justifiable*—but certainly not logically required identification of the world with the world for the conscious beings. But note: while this latter identification avoids the implausibility of idealism that results if it is wedded to a particular conscious perspective (producing the *solipsism* of the *I*-perspective,⁷ or the *anthropocentrism* of the *we*-perspective⁸), it brings to the fore another problem. This problem is already present, though not obvious, when we speak of our consciousness; it is, however, rather apparent when we speak of the consciousness of the conscious beings. Is there such a thing as the consciousness of the conscious beings?9 If idealism is to have chance, this expression, taken in some appropriate sense, must have a referent, since the world is certainly not in its entirety an object of the consciousness of each conscious being, or of the consciousness of *each* of us, or of *my* consciousness. It remains to be seen (see section 7) whether the problem of whether there is a *united* consciousness transcending the perspectives of particular subjects of consciousness, a consciousness that is all-encompassing in some appropriate sense, has a satisfactory positive solution.

It is already clear at this point that not every form of idealism is as plausible as every other. David Hume's idealism—an instance of atomistic idealistic panpsychism—is certainly an idealism of the less plausible sort. The objections against it are, at the same time, objections against the form of panpsychism it instantiates.

16.4 Why Atomistic Idealistic Panpsychism Is Unsatisfactory

Hume's peculiar idealism had a long aftermath, which is known as phenomenalism, a philosophical movement still very much alive even in the first decades of the twentieth century. There were several attempts to build a world from more or less atomistically conceived psychical elements, 'sense-data' as they came to be called. These attempts failed. Hume, in effect, knew that they would fail, although he himself does not speak of failure. What one cannot reconstruct on the basis of one's favored ontology (not even after serious effort) is usually considered an illusion (rather than an indicator of a defect in one's favored ontology). For Hume, then, both external objects and inner selves are ultimately *illusions*, inconsistencies with a semblance of existence. In Hume's view, Nature forces us, when we do not do philosophy, to accept what is, allegedly, rationally impossible: that external objects and inner selves exist and are persistent individuals and are (numerically) identically present *in their entirety* at every moment of their existence.¹⁰ The truth is that neither external objects nor inner selves can be satisfactorily reconstructed on the basis of a sense-data ontology. It is in accounting for these items that phenomenalism *fails* (though phenomenalists like Hume will not speak of failures that have to be admitted, but of illusions that have been, finally, revealed).

Thus, *atomistic* idealistic panpsychism has an insoluble composition problem. The aggregation and organization of sense-data, however complex, is not going to yield our familiar experiences of external objects or of ourselves, since it is already incapable of yielding what these experiences are *of*: external objects and ourselves.¹¹ The only way out is to *eliminate* (more properly speaking: to *deny the existence of*) these objects, with the accompanying *wholesale 'illusionizing'* of the experiences that seem to be directed at them. These are truly desperate measures. The better choice is to give up *atomistic* idealistic panpsychism.

16.5 Husserlian Idealism

Suppose the *basic* totality of actual being—*the world* in one sense of the word—consists of certain psychical events, of *(conscious) experiences* broadly speaking: perceptions, feelings, imaginings, remembrances, thoughts, volitions, and so on. All experiences belong to the basic totality of actual being, and *only* experiences belong to it.

This does not mean that everything is an experience.¹² Experiences usually have a bipolar structure: one of the two poles is the subject pole, the other is the object pole. At the subject pole is an experiencing subject, at the object pole is the object which is—or the entirety of the objects which are—being experienced in the experience.¹³ There may be experiences that lack an object pole and therefore an object. There is certainly no experience that lacks a subject pole; for all experiences have a subject, in fact (or so it seems), precisely one subject. If an experience has an object—let me define *the* object of an experience (for *all* cases where an experience has an object) as the entirety of the objects that are being experienced *in* it¹⁴—then all that that object *is in the experience* is found within the experience, is intrinsic to it. And yet the experience's object is not a part of the experience; for another experience (occurring perhaps much earlier or later) may have—and often *in fact* has—the very same object. Nor is the subject of an experience a part of it; for very different and temporally separated experiences may have—and often do have-the very same subject (literally the same: what is asserted is not that the subjects of the experiences are 'really' different, but temporal counterparts of each other, so that one can say in a sense that they are 'the same').¹⁵ Both the object and the subject of an experience are intrinsically determined by the experience (in other words, they intrinsically supervene on it). They are inseparable from it in the sense that the experience cannot ('cannot' taken in the strictest sense) exist without them as *that* experience. But, to repeat, they are not parts of it.

Every actual entity is an experience, or a subject of experience (that is, a subject of some experience), or an object of experience (that is, an object of some experience).¹⁶ At this point, and in the light of the previous two paragraphs, it should be clear that the metaphysical view presented in this section is (a) a form of ontological idealism (and hence of panpsychism, since every ontological idealism is *ipso facto* a form of panpsychism), and (b) that it promises to avoid the central shortcoming of Humeian idealism and phenomenalism. There seems to be room for external objects and inner selves in this other idealistic view-the inspiration of which is Husserlian phenomenology ('Phenomenology,' in short), as should be apparent to everyone who has come across that area of philosophy.¹⁷ This other view can be characterized as a form of idealism (therefore, of panpsychism), which pays due respect to experiential subjectivity and experiential intentionality (usually called 'phenomenal intentionality'), the intentionality of (and in) experience.¹⁸ In contrast, the fragmentation of experience into tiny bits, into sensedata, which we find in *atomistic* forms of idealism, leads to the irretrievable loss of the subject-centeredness and object-directedness of experience. And therewith—presupposing idealism—it leads to the irretrievable loss of inner

selves and external objects. This is so because the atomistic fragmentation of experience annihilates the polar, usually bipolar, structure of experience(s), which is a universal trait of all experience(s). (No wonder Hume was blind to intentionality.)

But so far there is only *the promise* that the Humeian nihilism about inner selves and external objects can be avoided in an idealistic ontological setting that is different from Hume's.¹⁹ On the way to showing that the view I have begun to describe does not just promise but also delivers, a crucial distinction must be made between what the object of an experience *is in the experience*, and what an object of an experience *is* (simpliciter). The distinction is familiar: it is the distinction between *appearance* and *reality* (more precisely, between the appearance of a thing, and the reality of *it*). In an idealistic setting, this distinction cannot be made in terms of mental representation and misrepresentation, for, in idealism, *all* of objective reality is 'in here,' rather than 'out there' to be represented or misrepresented 'in here.' The distinction must instead be made in terms of coherence and incoherence, harmony and disharmony, in ever widening experiential contexts, ever lengthening stretches of experience.

Thus, what an object of experience is is identical to what it is in the totality of all experiences, and is not necessarily (in fact, is usually not) identical to what it is in *this* particular experience (of which it is an object), or in *this* other particular experience. Let X be an object of experience. In the same way that the local context of an experience of which X is an object intrinsically determines what X is in that experience, so the global context of the totality of all experiences intrinsically determines what X is in that totality, that is, what X is (simpliciter). What X is locally (in an experience) may differ considerably from what X is globally (in the totality of all experiences). This is the Husserlian idealistic rendering of the distinction between appearance and reality. And if we ask ourselves how we actually make this distinction (we all make it), then it turns out that we do not do so in the way that the official representationalist epistemology of realism would seem to require. We do not do so by comparing a mental representation of X, and of how X is, with X itself, and with how X really is. (In fact, if mental representation were the basic cognitive relation we have to X, then X itself and how it really is would have to be inaccessible to us; we could then access only the mental representation of X, and how X is only as X is mentally represented to be, not as it really is in itself.) Rather, we put the appearances of X in the wider, and ever widening, context of our further experiences, which procedure is in perfect harmony with the epistemology of Husserlian idealism.²⁰ An object of experience that remains *stable* in this wider context emerges, by and by, as an objective reality with such-and-such objective properties.²¹

16.6 A Difficulty for Husserlian Idealism

There is a difficulty with the Husserlian idealistic way of distinguishing between appearance and (objective) reality: Different experiences may have—and often have in fact—the same subject of experience, but it certainly does not appear to be the case that all experiences have the same subject of experience. After all, there are my experiences and there are your experiences. I am not the subject of your experiences (or so it seems to me), and you are not the subject of my experiences (or so it seems to you). But if not all experiences have the same subject of experience, then the totality of all experiences is not itself an experience—but only a collection of experiences. This seems not only to contradict the Husserlian idealistic tenet (see the previous section) that every actual entity is an experience, or a subject of experience, or an object of experience,²² it *also* makes it difficult to conceive of contextualization-described in the previous section—as the 'maker of objective reality.' For a plurality of subjects of experience seems to render contextualization unfeasible. As long as experiences have the same subject, the contextualization of experiences has to straddle merely temporal separation. Over time, the experiences of one and the same subject come together in a fairly comprehensive totality which is itself an experience: the subject's stream of consciousness, which will contain many ultimate determinations—that is, determinations that are ultimate *for the subject*—of what an object is as opposed to what it seems to be (or, in other words, is in a particular, *local experience* of the subject). But the contextualization of experiences—which aims at the determination of objective reality (or, put differently, objective truth) in the form of stable content—cannot stop here. It must become *intersubjective*, must go from one subject's stream of consciousness to another's. And precisely this seems impossible. Each subject's stream of consciousness appears to be a closed world in itself (one feels reminded of Leibniz's 'windowless monads'). It is true that other subjects of experience and their experiences figure as objects, in, for example, my experiences; but they do so only in an indirect and—in comparison to the access I enjoy to myself and my experiences—in a very impoverished manner. (There is no direct perceiving by me of other subjects and their experiences, let alone of how they experience *me and my experiences*.)

My experiences (or *yours*) are an utterly insufficient basis for obtaining *objective reality* from experience (we are all convinced of this; this is why solipsism seems so ridiculous). What is needed is the totality of *all* experiences, not just the totality of the experiences of this or that subject. But, unfortunately, the totality of all experiences seems to be of no help either. For that totality is like a huge reservoir, which cannot be tapped—*unless* it were the totality of the experiences of *one* subject. Is there no way out? Ignore for a moment the assumption we all endorse (if we are in our right minds) and *suppose* that there is *just one* subject of experience. If there is just one subject of experience, then all experiences have the same subject of experience. The immediate benefit of this supposal is that the totality of all experiences is itself an experience—an experience whose object is itself a totality, namely, the whole of objective reality. Thus, the experience which is the totality of all experiences, *the world* in one sense of the word, has the whole of objective reality, *the world* in another sense of the word, as its object of experience (as its "intentional correlate," as Husserl would say). So far, so good. But *who* can believe that there is just one subject of experience? Doubtless, I would have to be identical to that subject, since, doubtless, I am the subject of experience of some experiences (*mine*); and you would have to be identical to it, too, since you too are the subject of experience of some experiences (*yours*). And therefore we are identical to each other—which is absurd. Or should it be the case after all that the only subject of experience, the subject of experiences, is *I*?

There must be a better way out than accepting solipsism. Let us recapitulate. Above I distinguished between what an object of experience *is in an experience* (of which it is an object) and what that object *is*, which latter phrase Husserlian idealism takes to mean the following: what that object *is in the totality of all experiences*. A more explicit way of making—in accordance with Husserlian idealism—the same distinction is to say that one must distinguish between what an object of experience *is in an experience for the subject of that experience,* and what that object *is,* which latter phrase Husserlian idealism takes to mean: what that object *is in the totality of all experiences for the subject of that totality.* This implies that the totality of all experiences is itself an experience and that it has a subject. But it prejudges nothing about the way in which *that* subject is related to *me* (or you).

16.7 A Solution to the Difficulty

The situation described in the last paragraph of the previous section invites the following manner of theorizing. Let us give up the proposition that every experience has *precisely one* subject, though certainly not the proposition that every experience has *a* subject. Let us accept, instead, the proposition that every experience has, intrinsic to it, precisely one *transcendental subject* (but *not* necessarily precisely one *subject*). In fact, every experience has one and the same transcendental subject, which Husserlian idealists simply call, aptly enough, *the* transcendental subject. Thus, there is no obstacle to assuming that the totality of all experiences is an experience (i.e., an experience of the transcendental subject, as are all experiences). The transcendental subject is not only the transcendental subject of that total experience, but also *the subject (tout court)* of it. Other experiences, however, not only have precisely one transcendental subject; they also have, in addition to it, precisely one *manifest* subject (also intrinsic to them), so that they have *two* subjects in total and one cannot speak of *the* subject of the experience. In all of *my* experiences, the transcendental subject (which is also the transcendental subject of *my* experiences. In all of *your* experiences, the transcendental subject of your experiences, the transcendental subject of your experiences. In all of *your* experiences, the transcendental subject of your experiences. In all of *your* experiences, the transcendental subject of your experiences again has a local projection, which is *you*, the manifest subject of *your* experiences. The manifest subject of your experiences and the transcendental subject of mine are the same.²³ This latter fact unites the two experiences into one experience—an experience belonging not to me or you, but to the transcendental subject.

These steps of theorizing are not mere arbitrary tricks for fabricating a way out of the difficulty. The transcendental subject is not a theoretical invention, a mere postulate; it has a place in the phenomenological description of experience. Some-though certainly not all-manifest subjects of experience can approach, and to a certain degree even approximate, the transcendental subject of their experiences, that is, the transcendental subject (tout court). They are capable of distancing themselves in a certain manner from their immediate individual experiences. And to the extent that they actually distance themselves from them, which is to say: put their experiences in context, are critical of them, correct them in the pursuit of *objectivity* (taken to be the synchronic and diachronic coherency, harmony, unity of all experiences with respect to objects of experience)-to that extent the manifest subjects of experience approach the transcendental subject, draw nearer to its point of view, perhaps come even near to it, in which case the transcendental subject is being approximated. In other words, there is in the experience of such manifest subjectsphenomenally manifest in their streams of consciousness—such a thing as the proto-objective point of view. The proto-objective point of view does not afford 'the view from nowhere' (to use the title of Thomas Nagel's famous book); the view from nowhere is an impossibility (not even the point of view of the transcendental subject affords a view from nowhere). The proto-objective point of view is still subjective, it is still the point of view of a manifest subject of experience, but of a manifest subject that distances itself from itself (so to speak) and approaches the transcendental subject. This self-distancing in approaching the transcendental subject is conspicuous in a question which we all feel has something to it but which—understood entirely literally, with no implicit orientation away from myself toward the transcendental subject—is just absurd (as

absurd as the question why 0 is 0): Why am I *I* (or *me*)? There is contingency in the projections of the transcendental subject, and a projection of it (i.e., a particular manifest subject), looking towards *its projector* and putting itself to some extent in its place (without necessarily being fully aware that it is doing this), may well ask itself for the reason for the fact that it is *this* projection.

16.8 Conclusions

None of the usual accusations against idealism (see section 3) touches Husserlian idealism. Husserlian idealism does not deny the existence of the physical (after all, many objects of experience are physical, whether in a straightforward or a scientific sense). Husserlian idealism does not propose that reality depends in all its aspects on the human will (although we have some influence on what we experience, we cannot just experience what we want: most of our experiences are neither wanted nor unwanted by us before they occur). Husserlian idealism avoids solipsism (I am not the only subject of experience, and not every object of experience is an object of some experience of mine). Husserlian idealism does not contradict the testimony of our senses since it accepts—as *objects of experience*—all objects that our senses present to us, and just as they are presented there. (But Husserlian idealism does not propose to view the objects of our senses in just one way: the way of simply taking them as such. My experiences of X at other times, my experience of X with a greater content and temporal extension than my original experience of X, experiences of X had by other manifest subjects—all of these experiences are bound to lead to corrected and, going in the direction of the totality of experience, increasingly objectively correct views of X.) Moreover, Husserlian idealism does not in the least contradict our natural empirical realism or its outgrowth, scientific realism. It only contradicts metaphysical realism (which some philosophers, unfortunately, find difficult to distinguish from scientific realism).²⁴ Husserlian idealism is entirely compatible with science, since the world that science presents to us—the objects, the properties, the laws—is the objective correlate, as Husserl would say, of a certain way of evaluating our experience, our consciousness, and thus remains entirely within the realm of (the objects of) experience.

16.9 The Big Question

Is Husserlian idealism also compatible with *religion*? I say it is. But in this case the compatibility is least obvious. Husserlian idealism is a form of *holistic*

idealistic panpsychism. The *holistic* aspect comes out not only in the fact that Husserlian idealism treats (intentionality-featuring) experiences as integrated bipolar wholes (with manifest and transcendental subject at one pole, and the object(s) of experience at the other, with all *the ways of being given*— a Husserlian expression²⁵—in between);²⁶ it also, and more conspicuously, comes out in the idea that all experiences combine into *one* big experience, with *one* subject: the transcendental subject. It seems natural to identify the transcendental subject with God.

The immediate consequence of this idea is that everything is in God (qua being in His total experience which, at the same time, is the totality of all experiences), whether as an experience, as a subject of experience, or as an object of experience. Note that the view that everything is in God is a theistic metaphysical position, which is known as *panentheism*. In contrast to pantheism, panentheism can be developed in such a way as to be compatible with traditional Christian theism.²⁷ Above (in section 7), the manifest subjects of experience were described as local projections of the transcendental subject; nothing precludes conceiving of all manifest subjects of experience—along with all experiences and their objects—as creations of the transcendental subject; that is, as created, limited simulacra of that subject. Nothing precludes giving the manifest subjects of experience, once they are created, some power of choice regarding which possibilities of experience in their range will become actual experiences (i.e., will be created). The idealistic perspective takes away nothing from the terrible seriousness of existence (which, if felt, brings forth the desire for *salvation*). On the contrary, it adds something to that seriousness: a real sense of the inner, the utterly intimate omnipresence of God.

16.10 A Solution to the Mind-Body Problem

From the standpoint of pure reason and experience—putting aside metaphysical preconceptions, as, for example, that only a materialistic view can be true—an explanation of the course of physical events in terms of the course of experiences is at least as rational as an explanation of the course of experiences by the course of physical events. Unfazed, the modern philosophy of mind looks solely in the direction of the latter explanatory approach (which is the inverse of the former). From the Husserlian idealistic point of view, however, no complete explanation of the course of experiences by the course of physical events is likely to be forthcoming, since, according to Husserlian idealism, physical events already depend on experiences for their *being and being-so*²⁸; the reason is that physical events are, if anything, *objects of experience*. The conspicuous failure of all extant attempts to produce a complete explanation of the course of experiences by the course of physical events suggests that no such attempt will succeed. A tentative reversal of the preferred explanatory direction seems worthwhile (if I may make a modest suggestion).

A reversal of this direction is mandatory on the fundamental level if Husserlian idealism is adopted as the basic metaphysical stance. Consider, then, explaining—under the presupposition of Husserlian idealism, on the fundamental level—the course of physical events by the course of experiences. It is important to bear in mind that this explaining is something entirely different from what is practiced, almost unthinkingly, in all empirical sciences, which is this: the *justifying* of hypotheses about physical events on the basis of our experiences and, so-called, inferences to the best explanation. For the aim of a Husserlian explanation is not to come to accept—as rationally as possible certain hypotheses about physical events because those hypotheses contribute (or seem to contribute) to a best *causal* explanation of our experiences in terms of physical events; its aim is to fundamentally explain the physical events in terms of experiences. In pursuing this latter aim, there is—under the presupposition of Husserlian idealism—no role whatsoever for causation to play (whereas causation plays an implicit but central role in the justificatory inference-to-the-best-explanation procedure just described; and *whereas* causation must play an *explicit* and central role if the goal is to *explain* the course of experiences by the course of physical events). In Husserlian idealism, and in the *fundamental* explanation of physical events by experiences,²⁹ the all-important hub is not causation but *intentionality*.

Doing without causation, at least on the fundamental level, is liberation. This is seen most conspicuously in the fundamental explanation of the mind-body relation. The wherewithal of this explanation is constituted by the intrinsic contents of our experiences of material objects, our experiences being taken separately and also in comparison, in other words: in their contextual juxtaposition to each other. The vehicle of the explanation of the mind-body relation is intrinsic determination (see section 5): my body intrinsically supervenes on my experiences. So do all other material objects I experience. However, in the case of my body, the intrinsic determination via my experiences is special. An object of experience is *my* body because, in addition to being experienced by others and me in the ways in which other material objects are experienced by others and me, it is also an (*intrinsic*) object of my experiences in a *special* way, in which it is not the object of the experiences of other subjects. That special way has many characteristics, but the three most salient ones are the following: (i) intimate experiential nearness of the object, (ii) immediate experiential control of the object, (iii) matching tactile experiences of (regions of) the object.

If I have the tactile experience of touching a region on the surface of the material object I call 'my body,' then I simultaneously have the tactile experience of being touched in precisely that region (and if the touch is accompanied by pressure, then the experience also reaches *below* that surface-region, goes deeper down); with other bodies, I do not have these matching experiences. In fact, I can outwardly delimit *my* body as that body which is the substratum of *matching* (in the sense just described) tactile experiences of *mine*. Note that a part of my body begins to feel to me just like any other material object (with *that* particular type of surface) if, with respect to that part, the matching of my tactile experiences fails to occur—say, because of local anesthesia, or because my arm has gone to sleep. And that feeling, that experience of otherness, is greatly consolidated if I experience—say, because my arm has gone to sleep—the absence of immediate control of that part in addition to the absence of matching tactile experiences with respect to that part. I can, at a given time, identify some (not all) prominent *parts of my body* as being those material objects which I experience at that time to be in my immediate control. The experience of immediate control is characterized by the experiential fusion of willing and *fulfillment*: neither time, nor effort, nor instruments interpose, in the experience, between the two.

The most important aspect of the *my-body* experience (which, in its totality, is what makes a body *my* body) is the experience—mainly nonvisual, to a lesser degree also visual—of the intimate nearness of the object. The subject experiences itself to be *in* a material object, to be *housed* by it—by a material object which is such that there is no other material object that also houses the subject and is nearer to the subject. Note that within the compass of my body—this region of inner warmth—the nearness has degrees: my hands I experience as being nearer to me than my feet, my eyelids as being nearer to me than my nose, my tongue as being nearer to me than my eyelids, etc. But none of these parts that I experience as mine do I experience as *housing* me. I am housed by the whole thing that those parts belong to, the region around the stomach being like the hearth of the house, exuding warmth.

16.11 An Objection

What is offered in section 10 is a mere sketch of how Husserlian idealism—a form of *intentionalistic* idealism and of holistic idealistic panpsychism—proposes to solve the mind-body problem *without recourse to any notion of causation*. However, enough was said to elicit an objection. What about the (living) brain and the entire (living) nervous system? In Husserlian idealism, they appear to be *pointless*—do they not?—in total neglect of their true importance. The answer to this objection is this: brain and nervous system, too, are objects of experience, but, typically, *not* of experiences of the subject to whose body

they belong, and, typically, only in an indirect way objects of the experiences of other subjects. The claim that brain and nervous system belong to a subject's body is indirect, and not a claim that is justified already on the *fundamental* level: beside the indispensable (but in itself insufficient) fact of bodily attachment (which also obtains for *hairs and nails*, and is, in that case, rather easily ascertained), their claim to belong to a subject's body rests on the fact that experience, if systematically explored, reveals that brain and nervous system are far from being pointless—in a very comprehensive and detailed way *causally relevant and vital* for a subject's consciousness and, indeed, for its existence as a manifest subject.

Nothing in this contradicts Husserlian idealism. In fact, experience reveals that many material objects (i.e., material *objects of experience*) are causally relevant and vital to the consciousness and existence of a manifest subject; the obvious example is the surrounding air (i.e., the by and large well-proportioned gaseous mixture of oxygen, nitrogen, and carbon dioxide).³⁰ This is simply the contingent way the world is; if the world is conceived of in an idealistic manner, that contingent way has not changed. Why should it?

Note, finally, that the demand that idealism ought to make some sort of deep sense of the contingent dependencies (of existence)—of us on our brains and hearts, on the surrounding air, on eating and drinking, on the right temperature, and so forth—is unjustified. If idealism does not make more sense of these dependencies than other views, then this does not count against it. (Note that the presently considered objection is different from the one in the first paragraph of this section.) Idealism ought not to be expected to eliminate *facta bruta*—because no metaphysical stance can.

Notes

- Neutral monism is traditionally connected with Spinoza. This is, strictly speaking, false. The nature of the *one substance* Spinoza postulates—of the *deus sive natura*—is *not* neutral; rather, it is *all-encompassing*, the physical and the psychical being just two of its infinitely many dimensions (or "attributes," as Spinoza says).
- 2. Modern panpsychism is an invention of the nineteenth century. An early criticism of it can be found in James (1890/1950). In chapter 6 of *The Principles of Psychology*, William James attacks what he polemically calls "the mind-stuff theory" (he also speaks of "mind-dust"); his description leaves no doubt that it is *atomistic dualistic panpsychism* which he attacks, although he does not use this designation. James nicely sums up the motivation behind modern panpsychism (he calls it "atomistic hylozoism"; James 1890/1950, 149), which is precisely the inability to accept (discontinuous, "jumpy") *emergence*:

If evolution is to work smoothly, consciousness in some shape must have been present at the very origin of things. . . . Each atom of the nebula. . . must have had an aboriginal atom of consciousness linked with it; and, just as the material atoms have formed bodies and brains by massing themselves together, so the mental atoms, by an analogous process of aggregation, have fused into those larger consciousnesses which we know in ourselves and suppose to exist in our fellow-animals.... [T]here must be an infinite number of degrees of consciousness, following the degrees of complication and aggregation of the primordial mind-dust. (James 1890/1950, 149–50.)

- 3. As William James fittingly remarks: "The fact is that discontinuity comes in if a new nature comes in at all.... The girl in 'Midshipman Easy' could not excuse the illegitimacy of her child by saying, 'it was a very small one'" (James 1890/1950, 149).
- 4. If one rejects dualism, then there are three basic (but further differentiable) options (if one wants to position oneself at all): physicalism, neutral monism, and idealism. For a panpsychist in the proper sense of the word (that is, if the expression 'psychist' in 'panpsychist' is to be taken seriously), physicalism and neutral monism are out of the question; for the psychical (in the proper sense of the word) is not physical, and it is not something neutral between the physical and the psychical, either.
- 5. Panpsychism, on the other hand, can very well be *not* idealistic (as we have seen). The first who spoke of 'idealistic panpsychism' appears to have been William James, using the word for a, broadly speaking, monadological metaphysics (Lamberth 1997, 249). But the designation was not an appropriate one: Lamberth notes that the position that James called "idealistic panpsychism"—and which Lamberth calls "strong panpsychism"—"retains a fundamental mind/matter dualism" (Lamberth 1997, 249). Curiously, James does not use the word "panpsychism" in the chapter of *The Principles of Psychology* (see endnote 2) which plainly addresses what modern thinkers consider to be *panpsychism* (Lamberth does not mention the chapter). Curiously too (in view of the criticism voiced in that chapter), James seems to have accepted a form of panpsychism after all, and *not* an idealistic one (see Lamberth 1997, 248–53).
- 6. "The limits of my language mean the limits of my world" (Tractatus, 5.6).
- 7. That is, the world is the world *for me*.
- 8. That is, the world is the world for us.
- 9. Is there even such a thing as *our* consciousness? There is a correlative problem: it is far from obvious that the singular terms 'the world for us,' 'the world for the conscious beings' designate anything.
- 10. On this matter see Meixner (2002).
- 11. The only thing that saves atomistic *dualistic* panpsychism from the same catastrophe seems to be its inherent lack of clarity. The psychical specks that (allegedly) sit on the atoms that compose a nervous system are not sense data. It is not known what they are; but as long as it is not known what they are, there is room for the hope that the aggregation and organization—i.e., the composition—of the atoms-*cum*-psyche will yield, ipso facto, the experiences of external objects and of ourselves.
- 12. Consider an analogy: Wittgenstein defined the world as the totality of *facts* (cf. the beginning of the *Tractatus*); it does not follow—nor did Wittgenstein assume—that everything is a fact, i.e., an obtaining state of affairs.
- 13. *In a sense*, also the experience is being experienced but not *in* the experience. Either it is experienced in another (so-called reflexive) experience, or experiencing an experience simply means *having it*.
- 14. Thus, even if an experience has several objects, one can speak of *the* object of the experience. One must only keep in mind that *the* object of an experience may consist of several objects of it. (This is the way the phrase 'the object of experience X' is understood here.)
- 15. On the idea of temporal counterparts see Meixner (2002).
- 16. Note that the 'or' in this statement is not an *exclusive* 'or' (an 'either...or...'). In fact, many experiences and subjects of experience are also objects of experience. Most objects of experience are, however, neither experiences nor subjects of experience. And a subject of experience is *never* (in fact, *cannot* ever be) an experience (and vice versa).
- 17. A good introduction is provided by Zahavi (2003) and Smith (2003). I speak of *Husserlian*, not of *Husserl's*, idealism, since the idealistic position I develop departs in certain respects from Husserl's. Husserl attempted to fulfill the exigencies of phenomenological foundationalism *and* to steer clear of solipsism (see Husserl's *Cartesian Meditations*). I neither

believe that accepting phenomenological foundationalism is a good idea, nor that solipsism can be entirely avoided if one does proceed on the basis of phenomenological foundationalism. Phenomenological foundationalism forces one to consider *the other (person)* as a sort of outgrowth of oneself, as being constituted within oneself. But the other person *is not* an outgrowth of oneself. For a detailed criticism along these lines cf. Meixner (2003); concerning Husserl's idealism see Meixner (2010).

- 18. On Husserlian intentionality see Meixner (2006) and Meixner (2010).
- 19. I am *not* saying that everybody wants to avoid these nihilisms. On the contrary, nihilism regarding inner selves is downright *fashionable* these days (and has been attractive for many philosophers for a long time).
- 20. Clearly, the rationale of this *putting into context* is, again, *comparison* for the purpose of determining what is objectively real and what is not. But it is a kind of comparison that is very different from the kind of comparison that has just been described (in the main text): it does not cross the borders of immanence.
- 21. *Representationalist realists* are confronted with the absurd situation that they cannot help practicing contextualization, although, from their point of view, it is far from obvious *why* contextualization should help us find out about objective reality.
- 22. However, one can avoid this particular problem by stipulating that, by the word 'entity,' one always means a single, at least minimally unified item, not a plurality or collection.
- 23. It is instructive to reread sections 5 and 6 in the light of interpreting 'the subject of' either as 'the *transcendental* subject of' or as 'the *manifest* subject of.' Both interpretations are (knowingly) conflated in those sections.
- 24. It is sometimes argued that the undeniable success of science requires for its explanation the truth of realism. But if the success of science requires for its explanation the truth of realism, then it cannot be the truth of *metaphysical* realism. For it is not required for the success of science (let alone for the explanation of that success) that there be objects that are not objects of any experience.
- 25. Another Husserlian expression is rather more familiar: *intentionality*. The many ways of being given are the many ways of *intentionality*.
- 26. It goes without saying (but I say it nevertheless): the spatial metaphors in the parenthesis that this endnote refers to must not be taken literally.
- 27. I am referring primarily to Christian theism since this is the only theism of which I have sufficient knowledge.
- 28. This is the English translation of the useful German phrase Sein und Sosein.
- 29. Husserlian idealism can accommodate the cognitive activities spoken about in parentheses in the previous sentence of the main text—and the idea of causation that is central to those cognitive activities, and the realism which those activities presuppose—as higherorder activities and constructions (i.e., as derivatives). However, in this accommodating, matters are no longer considered with respect to *the fundamental level* of explanation.
- 30. This example also shows that not even the inexorable necessity of a material item for the existence of a manifest subject is, in itself, sufficient for including that item in the subject's body.

Bibliography

James, William. The Principles of Psychology. Vol. 1. New York: Dover, 1890/1950.

- Lamberth, David. "Interpreting the Universe after a Social Analogy: Intimacy, Panpsychism, and a Finite God in a Pluralistic Universe." *The Cambridge Companion to William James*. Ed. Ruth Putnam. Cambridge: Cambridge University Press, 1997. 237–59.
- Meixner, Uwe (2002). "Change and Change-Ersatz." Individuals, Essence and Identity. Themes of Analytic Metaphysics. Ed. Andreas Bottani, Massimiliano Carrara and Pierdaniele Giaretta. Dordrecht: Kluwer, 2002. 427–49.

- Meixner, Uwe. "Die Aktualität Husserls für die moderne Philosophie des Geistes." *Seele, Denken, Bewusstsein.* Ed. Uwe Meixner and Albert Newen. Berlin: De Gruyter, 2003. 308–88.
- Meixner, Uwe. "Classical Intentionality." Erkenntnis 65.1 (2006): 25-45.
- Meixner, Uwe. "Husserls transzendentaler Idealismus als Supervenienzthese ein interner Realismus." *Husserl und die Philosophie des Geistes*. Ed. Manfred Frank and Niels Weidtmann. Berlin: Suhrkamp, 2010. 178–208.

De Spinoza, Benedictus. Ethica. Ordine Geometrico Demonstrata. Stuttgart: Reclam, 1977.

Smith, David. Husserl and the Cartesian Meditations. Milton Park: Routledge, 2003.

Zahavi, Dan. Husserl's Phenomenology. Princeton, NJ: Princeton University Press, 2003.

INDEX

actual occasion, 4, 51 agency, 48, 61-63, 69, 136, 144, 158 agent. See agency aggregates, 10, 115, 120-21, 175, 185-86, 242-43, 296, 306 agnosticism, 51, 53-54, 56, 371 Albantakis, Larissa, 67 Albert, David, 330 Alexander, Samuel, 107, 319, 330, 340 Allen-Hermanson, Sean, 1 Anscombe, Elizabeth, 372-73, 375-76, 382 Ansichsein, 7, 75, 80-81, 105 anthropocentrism, 391 antirealism, 53-54 argument anti-aggregation, 185, 242 causal, 20, 23-24, 27, 29-30, 32 conceivability, 20, 22-24, 27-28, 30, 32, 36-37, 43-45, 130, 143, 145, 187-88, 204, 243, 249, 290-91, 322, 365 epistemic, 22, 31 genetic, 3, 7, 12, 50, 68, 356-57 Hegelian, 6-7, 20, 30, 44-45 from intrisic natures, 3, 56, 68 knowledge, 22, 188, 243, 322, 325, 376, 381 model-theoretic, 54 palette, 9-10, 145-47, 189, 244 against physicalism, modal, 243 private language, 379 revelation, 9-10, 147, 149, 190, 245 structural mismatch, 10, 190-91, 202, 206, 210-11, 245 subject-summing, 10, 186, 197, 242-43, 256-58 zombie, 29, 51, 58, 143, 227, 269-70, 278 Aristotle, 1, 52, 76-77, 105 Armstrong, David M., 372, 375 atomism. 7 Audi, Paul, 330

Augustine, 306, 370 awareness of awareness (see awareness: second-order) first-order, 280 first-person, 380-81 first-person self- (see awareness: first-person) mental, 379 second-order, 105-7, 271-72 self-, 272, 376, 378, 381-82 Baars, Bernard, 167, 171-72 Baker, Lynne, 381 Barnes, Elizabeth, 301 Barnett, David, 41 Bartlett, 331 Basile, Pierfrancesco, 107, 211 Bayne, Tim, 276 Bealer, George, 372 Bedau, Mark, 132, 347 behaviorism, 108, 374 Bennett, Jonathan, 330 Bennett, Karen, 45, 133 Berkeley, George, 15, 88, 199 Blamauer, Michael, 15, 211 Block, Ned, 216-17, 225, 274 Bohm, David, 2-3, 107, 119, 276 Bohr, Niels, 101 Bolender, John, 46, 358 boundary problem, 155-56, 162, 183, 250, 253, 260, 266 brain states, 90, 130, 135, 240-41, 291, 319, 374, 382. See also neurophysiology bridging problem, 251-53, 261-63, 274-75 Broad, Charles Dunbar, 132-33, 232-33, 313, 318-20, 326, 329 Brogaard, Berit, 9, 138-39, 148 Bruno, Giordano, 48 Brüntrup, Godehard, 52-53, 334-35, 338-42, 344-45

Buck, Alexander, 128 Buckareff, Andrei, 128 bundle theory, 107, 364 Byrne, Alex, 148, 278 Campbell, Charles, 380 Carnap, Rudolf, 277 carrier, 55-56, 166-67, 169, 174. See also ultimate-carrier problem pyramid, 172 (see also synoptic pyramid) receptive, 167 Carroll, John, 135 causal closure, 29-30, 45, 48, 60, 126-27, 194, 295,374 causal efficacy, 14, 60, 63, 126, 245, 295. See also causation: mental causal exclusion, 29, 45, 66, 210 causal nexus, 65, 157, 161-63, 165-67, 169-70, 175 causal significance, 157, 159, 161-66 causation, 29, 76, 82, 135, 159, 166, 200, 374, 400-1 backward, 160 carrier-, 155 downward, 6, 48, 60, 65-67, 69, 193 final, 165-66 ineliminability of, 347 mental, 27, 39-40, 48, 57, 60, 62-64, 69-70, 161, 166, 181, 193-94, 199, 205, 245, 327, 328, 369, 371, 374-75, 380, 385 non-Humean, 158 psycho-physical, 246, 375 Cavendish, Margaret, 211 cellular automaton, 55, 58-59, 63, 66, 342 Chalmers, David J., 20-21, 27, 40, 60-61, 103, 114-25, 130-39, 143-47, 172, 179-87, 202-10, 216-17, 222-27, 236, 241-46, 249-63, 266-81, 288-89, 295, 319-20, 351-65, 383 Churchland, Paul, 279, 376, 382 circularity, 55 Clark, Andy, 342 Clarke, David, 356 Coleman, Sam, 41, 44, 180, 202, 125-223, 225, 321, 353, 355 Collins, Francis, 109, 384 combination laws of experiential, 100, 103 mysterianism about, 294, 323 combination problem grain, 103, 183, 190-91, 250, 253, 260, 267-69 palette, 103, 147, 183, 189, 204-5, 250-53, 260, 263 quality, 44, 182-90, 201-4, 210, 217, 250, 251-52 structural mismatch, 103, 183, 190-91, 195-99, 210, 250-53, 260-61, 269-71 structure, 44, 182-83, 188, 201-6, 210, 217

subject, 39, 103, 144-45, 182-84, 197-204, 210, 217, 249-54, 257, 260, 275 unity, 183, 250, 253-54, 260 composition restricted, 132-33, 296 special, 69 (see also special composition question) unrestricted, 296, 301 unrestricted phenomenal, 12, 296, 299 (see also composition: unrestricted) conceivability, 22-23, 29, 36-38, 43, 118, 187-88, 226-27, 269, 287-91 conceivability principle, 12, 288 transparency, 12, 289-91, 301 two-dimensional, 288-91, 301 conceptual isolation of subjects, 12, 287, 291, 293,300 consciousness cosmic, 8, 113, 116-17, 120-24, 156, 194, 196 evolution of, 4, 238, 241 o-, 295 phenomenal, 21, 58-59, 130, 230-31, 251, 254, 269, 306-7, 311, 313, 316, 318, 319-22, 325, 328-29, 334 proto-, 20, 31, 171, 389 (see also proto-mentality) contamination problem, 252-53, 260, 262 continuum hypothesis, 264-65, 276 Cook, Paul, 276 Corcoran, Kevin, 372-73, 375-76, 382 cosmopsychism, 20, 117, 194, 196 derivation problem for, 8, 116–17, 120–24, 127, 211 existence, 116-17 identity, 195-96 no sign problem for, 123, 126 no test problem for, 125-26 priority, 8, 113-14, 116-28 counterfactuals, 79, 157, 374 Crick, Francis, 139 Dainton, Barry, 211-12 Davidson, Donald, 60-61, 63, 69 decoherence theory, 2, 163 decomposition, 8, 113, 118-20, 127-28, 308-12, 315-16 problem, 127, 194, 196, 261 deflationism, 40, 42, 49, 68, 197-98, 202-5, 210, 271, 299 Dehmelt, Hans, 119 Dennett, Daniel, 231, 375-76, 382-85 Descartes, René, 49-52, 56, 68, 77-78, 103, 306, 353, 370, 378-81 determinism, 255, 295, 340 disposition, 26-27, 35, 107, 157, 181, 283-43, 360 Dominus, Susan, 220 Dorsey, Jonathan, 366 Dretske, Fred, 280, 311

dualism emergent, 7, 14, 60, 65, 67-70 property, 35, 320, 380 substance, 39, 48-49, 141, 193-94, 284-85, 356, 374 Eccles, Sir John, 141 Eddington, Sir Arthur, 2-3, 94, 101, 106 Edelman, Gerald, 167 Einstein, Albert, 101 eliminativism, 197-98, 202, 204, 334, 374, 376 emergence causal, 67, 274 conservative, 232-34, 236, 238, 242 diachronic, 335, 340-42 evolutionary, 13 interattribute, 50, 68, 334, 340-41, 345 intraattribute, 50, 68, 340-41, 345 of a nonreductively conceived mental realty, 3 ontological, 133, 326 property, 131, 310, 317-18 psycho-physical, 388 radical, 82-83, 91-99, 102-6, 109, 232-39, 242, 246, 352-53 strong, 25, 51-52, 59-60, 63-69, 115, 118, 120, 127, 130-35, 137, 141, 150, 154-55, 166, 181, 192-93, 335, 347 superstrong, 7, 68-69, 345-46 weak, 59, 66, 68, 131-32, 335, 340-41 endurants, 59 epiphenomenalism, 24, 30, 39-40, 161, 166, 193-97, 206, 246, 320 Epperson, Michael, 2 essentialism dispositional, 26, 290, 295, 300 scientific, 4-5, 290, 295, 300 Evans, Jil, 384-85 evolution of consciousness, 91, 141, 173 process of, 2, 225 theory of, 94, 107, 110 experiencer. See subject experientialism, subatomic-, 316 explanation, reductive, 59, 249, 314, 317, 334, 342-47 Faraday, Michael, 107 Feigl, Herbert, 41, 46, 203, 274, 277, 320, 330 Feynman, Richard, 101 Fichte, Johann Gottlieb, 44 field experiential, 253, 262, 266, 268, 275, 278-79 quantum, 108-9, 222-23 (see also quantum physics) visual, 42-43, 121, 191, 207-9, 220, 375 Fine, Kit, 44, 147, 301 Flanagan, Owen, 331 Flew, Antony, 371

Fodor, Jerry, 352 Foster, John, 109, 274, 277 Franklin, Sir Benjamin, 327 Frauenstädt, Julius, 107 free will, 141 Frege, Gottlob, 136, 230 functionalism, 23, 34, 202-4, 268, 277, 298, 356 fusion, 67, 100, 238-47, 264, 401 Gabora, Liane, 193, 211 gap causal, 24 explanatory, 22, 100, 204, 212, 216, 226, 230, 254-55, 326-29, 344 nonsubject/subject, 38, 41-42 ontological, 22, 31 quality/awareness, 43, 363-64 subject-subject, 37, 41-42 Geach, Peter, 381 Gennaro, Rocco, 279 global workspace, 9, 167-68, 171-72 Goecke, Benedikt Paul, 372 Goetz, Stuart, 379, 384 Goff, Philip, 36, 114, 180, 186, 190, 200, 215, 226-27, 242, 256-57 Goldstein, Shelly, 330 Goodale, Melvyn, 247, 305 graviton, 9, 137-38, 140-41 grounding, 21-22, 49-55, 65, 194, 353-54, 360, 364 Hacker, Peter, 371 Haldane, John, 306 Hameroff, Stuart, 2, 4, 241 hard problem, 130-31, 139, 150, 155, 320-21, 377, 383 Harman, Gilbert, 280 Hartshorne, Charles, 264, 276 Hassel Mørch, Hedda, 35, 107, 211 Haugeland, John, 55 Hawking, Stephen, 85 Hegel, Georg, Wilhelm Friedrich, 6-7, 20, 24-25, 27-30, 32, 40, 44-45 Heil, John, 35 Heisenberg, Werner, 76 Hellie, Benj, 83 Higgs, Peter, 76, 139-40 Hilbert, David, 148 Hilev, Basil, 3 Hill, Christopher, 311, 326, 331 Himma, Kenneth, 374-75 Hoel, Erik, 67 holism, 7, 196, 199, 210, 261. See also panpsychism: holistic Holman, Emmett, 5 Holt, Edwin B., 351 Homer, 95 homogeneity, 49-50, 120, 277

HOT, 11, 250-51, 265-70, 273-74, 276-81 Huebner, Bryce, 227 Hume, David, 93, 95, 158-59, 387, 391-94 Humphrey, Nicholas, 377-78 Humphreys, Paul, 67, 239, 247 Husserl, Edmund, 15, 390, 392-96, 398-404 Huxley, Thomas, 318 Hylal, 87-89, 90, 93, 100 hylomorphism, 52, 55 hylozoism, 402 idealism, 35, 48-49, 51, 54, 68, 70, 92, 94, 194-95, 350, 380, 382, 387, 390-95, 402 absolute, 1, 94 Husserlian, 392, 394-96, 398-402, 403-4 intentionalistic, 401 ontological, 390-91, 393 Russellian, 34-35, 46 identity, 35, 41, 57-58, 62, 75, 81, 102, 143, 154, 180, 219, 224, 236-37, 285, 327-29, 376, 380, 391 numerical, 265 personal, 41, 198, 376 a posteriori, 13 qualitative, 265 Russellian, 34-35, 45-46 theory of, 34, 36, 46 incommensurability problem, 252-53, 260, 263-65, 276-77 indeterminism, 65-66, 342 inference to the best explanation, 4, 389,400 information active, 2 flow of, 57 integrated, 67, 150, 168, 171, 175, 193, 212 (see also information: flow of) infusion, 11, 198-200, 212, 238-46 ingression, 164 intensions primary, 5-6, 46, 301 secondary, 5, 46, 301 intentionality, 136, 139-40, 172, 393-94, 399-400 interactionism, 24, 30, 39-40, 165-66, 193 introspection, 41, 147-50, 175, 190, 202, 218-21, 238, 245, 273, 293 intuition, zombie. See argument: zombie irreducibility of emergent properties, 13, 275, 278, 343-44, 346 Jackson, Frank, 108, 243, 347, 324, 327, 376 James, William, 36, 41-44, 50, 81, 103, 143-45,

James, William, 30, 41–44, 50, 81, 103, 143–48 154–55, 179–87, 197, 202–4, 216–17, 224–27, 236–38, 242, 254, 285–87, 291, 320–21, 349–55 Jaskolla, Ludwig, 16, 128

Kant, Immanuel, 51, 53-54, 56, 70, 77-80, 95, 103, 106-7, 217 Kenny, Sir Anthony, 371 Kim, Jaegwon, 65-66, 133, 343, 372-73, 375,383 knowledge, a priori, 243, 291 Koch, Christof, 3, 139 Koons, Robert, 372 Kriegel, Uriah, 110, 271-72, 276-77, 279-80 Kripke, Saul, 5, 22, 147, 247, 287-89, 300, 327-28 Kuhn, Thomas, 217, 227 Ladyman, James, 107-8 Landini, Gregory, 358, 366 Lavazza, Andrea, 372 laws of nature, 23, 25, 27, 29, 33, 60, 78, 130, 134-35, 141, 193, 232, 242-43, 320, 326, 339, 354, 362, 374, 388-89 Lehrer, Keith, 380 Leibniz, Gottfried Wilhelm, 40, 49, 51-53, 55-56, 63, 68, 76, 78, 81, 106, 195, 237-38, 370, 395 Levine, Joseph, 109, 230, 326, 343-44 Lewis, David, 78, 132, 145, 290, 300-1 life worlds, 59, 66 Loar, Brian, 230, 331 Locke, John, 370 Lockwood, Michael, 45-56, 103, 109, 183, 190, 241, 267-68, 276-79 Loewer, Barry, 330 Lovejoy, Arthur, 321 Lovering, Rob, 379-80 Lucretius, 109, 252, 254, 274 Lycan, William, 277 Mach, Ernst, 41, 197, 204, 349, 351, 354, 366

- Mary (Neuroscientist), 10, 144, 188-89, 243 materialism, 20-25, 27, 31, 36, 41, 44, 48, 83-84, 88, 90, 92-94, 101-4, 179, 182, 190, 344, 347, 356-57, 372, 375, 385, 399 type-A, 23, 31, 35 type-B, 4-5, 9, 23, 25, 31, 35 Mathews, Freya, 128 Maudlin, Tim, 135, 330 Maxwell, Grover, 34, 46, 183, 191 McFague, Sallie, 370 McGinn, Collin, 229, 306, 323-25, 331 McLaughlin, Brian P., 133, 247, 301, 318, 325-26, 330-31, 344 McWeeny, Jennifer, 211 Meixner, Uwe, 14-15, 403-4 mental content, necessary total overlap of, 220 menton, 9, 130, 137-38, 140-43, 145-46, 148,150 mereology, 119, 134, 145, 306, 310 Merker, Björn, 146-47
- Merricks, Trenton, 150, 301, 371

Horgan, Terence, 108, 128, 301

metaphysical isolation of subjects, 287, 291-94 microdeterminism, 59, 65 microexperiences, 24-25, 27, 29-30, 36-37, 40-41, 49, 57, 99-100, 131, 144, 180, 183-84, 188, 190-95, 197, 199, 202, 217, 245, 274, 351-53, 362 micropsychism, 103, 114 Mill, John Stuart, 234-36, 238, 246-47 Milner, David, 247, 305 mind dust, 179, 305-6, 313, 316, 329, 402-3 modality absolute necessity, 242 de re, 147 epistemic possibility, 5 metaphysical possibility, 5, 57, 89, 119, 290 a posteriori, 25, 31, 35, 57, 59, 288, 317, 327, 329 potentiality, 2, 159-65, 175, 389 a priori, 22, 25, 31-32, 35, 38, 41, 57, 59, 77, 79, 81, 119, 145, 182, 189, 243, 279, 288–91, 295, 298-99, 309, 317, 322, 325-28, 389 strong entailment, 22, 25, 31-32, 57, 182, 309, 328, 375, 385 model building block, 39, 64, 119-22, 138, 340, 352, 359 field, 138 monad, 40, 49, 51, 123, 195, 200, 237, 293, 395,403 monism dual-aspect, 7, 51 existence, 108, 115, 128 experiential-hylal, 88-92 idealistic, 51 materialistic, 83, 92 mental, 355 naturalistic, 102 neutral, 35, 41, 48, 51, 57, 70, 81, 91-92, 155, 197, 203-4, 249, 313-14, 349, 350-59, 363-65 panpsychist neutral, 51, 355-56 physical, 13, 334-36, 338, 355 priority, 8, 113-17, 122-23, 128, 356 Russellian, 32-44, 49-51, 157-59, 167, 203-11, 244, 295, 349, 353 Russel's neutral, 14, 51, 349, 353, 358, 363-66 stuff, 83, 88, 90, 98, 108 thing, 83, 86-87, 94 type-F, 5, 7, 128 Montero, Barbara Gail, 11, 46, 118, 128, 211, 227, 324 Moore, George, 90, 97, 375 More, Henry, 375 Morgan, Lloyd, 340 Nagasawa, Yujin, 8, 128, 366

Nagel, Thomas, 232–34, 244, 246, 305, 307–14, 319–20, 322–26, 356, 369, 376, 397

naturalism, 82-92, 102-4, 166-71, 229, 250, 319, 336, 372, 381-88. See also real naturalism necessity, brute, 23 neurobiologicalism, 13, 329 neurophysiology, 58, 99-100, 372, 374. See also neurobiologicalism Newman, Max, 53-54, 86 Newton, Sir Isaac, 327 Nida-Rümelin, Martine, 41 Nietzsche, Friedrich, 78, 82, 110 objection, structural mismatch. See argument: structural mismatch Occam, William of, 101-2, 104, 221-22, 369, 391 Occam's razor, 101-2, 104, 221-22, 369, 391 O'Connor, Timothy, 65-66, 68-69, 133-34, 136 organism, 130-45, 150, 172, 201-3, 265, 284-98, 306-13, 338-47, 356 overdetermination, 23-24, 30, 40, 45, 194 causal, 24, 134 panentheism, 399 panexperientialism, 19, 51, 80, 93, 167-71, 306-18, 320-26, 355, 365 panprotoexperientialism, 314, 317, 320, 322-26, 329-30 panprotopsychism, 19-20, 30-44, 131, 179-210, 249-50, 314, 320 panpsychism atomistic dualistic, 14, 388-89, 402-3 atomistic idealistic, 14, 388, 391-92 autonomous, 192, 194-95 combinatorial infusion, 198-200, 211 compositional, 51, 339, 344, 347 constitutive, 25, 29, 36-37, 40-41, 48-49, 52, 57-65, 69-70, 131, 137, 144-47, 149-50, 181, 184, 186-87, 190-92, 202, 206, 245-46, 295, 297, 299 deferential emergentist, 234 deflationist, 197-98, 202-5, 210 emergent, 68, 181, 192-94, 199-200, 211 holistic, 8, 195 holistic dualistic, 14, 387-88 holistic idealistic, 14-15, 387-88, 398-99, 401 identity, 192, 195-97, 199-200, 212 noncompositional, 345, 347 nonconstitutive, 25, 30, 48-51, 60-70, 181, 192 phenomenal bonding, 200-2, 210 (see also phenomenal bonding) Russellian, 26-33, 139-40, 181-99, 206-9, 244, 249, 254, 353-54 panqualityism, 40-44, 203-4, 249-70, 273, 354, 362-65

pansubjectivism, 307, 313-14 pantheism, 117, 127, 399 Papineau, David, 45, 230, 300-1, 344 Parfit, Derek, 41 Pauen, Michael, 343 Penrose, Roger, 2, 4, 101, 241 Pepper, Stephen C., 41, 274 Pereboom, Derk, 277 Perry, John, 81 person, 63, 128, 135, 144, 256, 334, 345, 365, 369-85 perspectival pooling, 217-21, 223, 226 perspective, 97, 161-62, 166, 175, 209, 218-21, 224-25, 255-57, 314, 335-37, 340, 344, 346, 363, 381-82, 391, 399. See also point ofview phenomenal bonding, 200-2, 210, 256-59, 277-78, 283, 292-93, 295-96, 300 relation, 21, 200-1, 224, 257-58, 292-94, 296, 299-300 phenomenal borderline case, 298-99 phenomenal character. See what it's like phenomenal concepts, 57-58, 230-31, 277, 291-92, 302 phenomenal experience, 56, 59, 114, 117, 120-21, 127, 141, 148-50, 231, 240, 346 phenomenal mind, 3-4, 50, 58-60, 65, 68-69 phenomenal states, 10, 36, 138, 182, 200-1, 257, 269, 342 phenomenal transparency, 12, 291-94, 299-301 phenomenalism, 94, 392-93 phenomenality, 69, 113, 117, 124-25, 127, 218 micro, 24-37, 57, 61-68, 130-33, 141-50, 181-209, 246, 295, 353 proto-, 31-35, 37-38, 42, 61, 113-17, 119-21, 124, 126-27, 130-31, 137, 139, 182, 184, 188-89, 191-92, 198, 203-4, 207 physicalism nonreductive, 3, 62, 130 reductive, 52, 130-37, 142-43, 224, 317-29, 334, 338-40 type-A, 70 type-B, 57-58 physicSalism, 52, 335-36, 338-40, 342, 344-47 Planck, Max, 4, 56, 316 Plantinga, Alvin, 376-77 Plato, 50, 107, 166, 306, 370-71 Platonism, 166 Poincaré, Henri, 108 point of view, 64, 219, 255-59, 325-29, 351, 362, 381, 397 possible worlds, 50, 58, 62-63, 87, 132, 142, 232, 242-43, 288, 309, 319, 327 Potrč, Matjaz, 108, 128, 301 Powell, Jonathan, 241 Price, Henry, 375 primacy of the mental, 14, 369, 371, 380

primordial feeling, 131, 139-40, 142-43, 145-47, 149-50, 369, 371, 380 Prinz, Jesse, 240-41, 244, 247 problem of other minds, 56 problem of the many, 243 production problem, 251-52, 260 properties categorical, 26-28, 35, 52, 77, 158, 181, 200, 246, 353, 360, 364 derivational, 309 determinable, 122-23, 138 determinate, 123, 328 dispositional, 26-27, 34-35, 182, 244, 353-54, 360, 364 effective, 162-67, 169-70, 174 intrinsic, 26, 51, 53-56, 58, 60-63, 68, 78-79, 85, 88, 92, 99-101, 132-33, 140, 145, 150, 153, 160, 166, 169, 172, 255-56, 258-59, 310-13, 318, 324, 338-39, 345, 361 nonsystemic, 336-37, 247 o-physical, 33, 45-46 receptive, 162-66, 170 second-order, 27 structural, 25-27, 31-32, 34, 37, 42, 51-55, 57-59, 63, 68, 86, 89-90, 134, 145, 200, 207, 239, 272, 310-13, 316, 318-21, 323-24, 326-27, 329, 346, 353-54, 361, 364 systemic, 13, 334-39, 343 t-physical, 33-35, 45-46 Prosser, Simon, 67 proto-experience, 37, 41, 91, 93, 184, 197. See also qualia: protoproto-mentality, 3-4, 11, 13, 34, 58-59, 322, 335-37, 339-41, 344, 346 psychology, 55, 138, 157 psychon, 141 Putnam, Hillary, 5, 54 qualia, 56, 230-31, 249, 253, 264-65, 271, 307-9, 311-26, 328-29, 355 macrophenomenal, 141, 146-47, 189, 366 phenomenal, 56, 146, 183, 190, 204-5, 264, 293, 336-37, 344, 355 proto-, 314-15, 317, 323-26 unexperienced, 11, 249, 274, 335 qualitativity, 75, 78-79, 89-90, 92, 97, 104, 264 quantum physics, 67, 86, 96, 100, 102-3, 140-42, 162-63, 196, 199, 210, 222-23, 238-39, 241, 245, 247, 261, 275-76, 284 quiddities, 26-42, 51-52, 59-62, 68, 181-84, 196, 200-9, 338, 353-54, 360-64 Quine, Willard van Orman, 95

Ramsey sentence, 27 real naturalism, 83–85 realism epistemic structural, 53–54, 97

ontic structural, 86 scientific, 53-54, 398 receptive binding, 164 receptive connection, 163-67, 170, 172-73 receptive connectivity, 169 receptive context, 165 receptive nexus, 174 Regius, Henry de Roy, 108 relation, binding, 161-62, 165, 169 representation, 136-38, 141, 150, 160, 171, 173, 208, 239-40, 265-67, 269, 305, 394 res cogitans, 335 Rescher, Nicholas, 4 Rey, Georges, 230 Robinson, Howard, 109, 372 Roelofs, Luke, 211-12, 301 Rosenberg, Gregg, 46, 55, 58, 108, 161, 171, 180, 183, 193, 200, 330 Rosenthal, David, 250-51, 265, 267, 274, 276-77, 280 Ross, Don, 108 Russell, Bertrand, 26, 41, 51, 53-54, 56, 68, 80-81, 91, 96-97, 106, 136, 197, 204, 349, 351, 353-54, 358-66, 381 Russellianism, 14, 253, 355, 358, 361, 366 Ryle, Gilbert, 371 Salerno, Joe, 148 Sayre, Kenneth, 350 Schaffer, Jonathan, 44, 67, 108, 115-16, 118-19, 122-23 Schlick, Moritz, 76, 381 Schopenhauer, Arthur, 78, 86, 106-7 Schrödinger, Erwin, 3, 101 Schrödinger's equation, 3 Seager, William, 36, 108-9, 120, 143, 180, 198-99, 211, 242, 247, 274, 365-66 Searle, John, 9, 125, 138, 229, 330 Seeley, Thomas, 225 self, 59, 97, 104-6, 124, 167, 181, 229, 247, 255, 272, 305, 359, 362, 370, 376-78, 380-85, 397 -intimation, 105-6, 148 Sellars, Wilfrid, 275, 277 semantics, 33, 35, 45, 58 sensation, 80-81, 235, 265, 267, 346, 356, 359 sense-data, 53-54, 208, 305, 358-59, 366, 392 Shani, Itay, 211, 216 Shoemaker, Sidney, 330 Sider, Theodore, 145, 150, 300-1 Siewert, Charles, 247 Skrbina, David, 1, 15, 211, 307, 349, 351, 355-56 Smart, John, 34, 330 Smith, David, 403 Smolin, Lee, 276 solipsism, 15, 374, 390-91, 395-96, 398, 403-4 Sosa, Ernest, 61, 69-70

soul, 20, 49, 103, 197, 223, 226, 306, 370-71 spacetime, 33, 49-51, 55, 58, 68, 76-77, 81, 83, 85-86, 89-90, 95-97, 114, 121, 123, 132, 138, 156-57, 159, 191, 208, 217, 264, 267, 286, 306, 312, 359-60, 372, 374-75 special composition question, 12, 64, 296 Spinoza, Baruch, 78, 81, 106, 108, 387-88, 402 split brain, 219-21, 227 Sprigge, Timothy, 99, 305, 376 Stapp, Henry, 2, 241 Stephan, Achim, 13, 15, 340-44, 347 Stoljar, Daniel, 33, 44-46, 109, 191, 208, 215-17 Strawson, Galen, 33, 44-46, 90, 103, 107-10, 115, 211, 247, 275, 334, 347, 356-57, 371, 384 stream of consciousness, 231, 395 string theory, 263, 276 structuralism, epistemic, 53 Stubenberg, Leopold, 13-14, 354, 358 Sturgeon, Scott, 331 subject, 144 conscious, 21, 36-37, 42, 146, 181, 186, 188, 195, 242, 285, 295, 297-99, 382 of experience, 39, 41-43, 64-65, 80-81, 94, 103, 169, 182, 184, 195, 197-98, 217, 254, 286, 292-93, 305-6, 314, 416-17, 351, 354, 362-63, 393, 395-99, 403 experiences without, 42 experiencing, 144, 156-57, 212, 393 transcendental, 396-99, 404 subjectivity, 51, 64, 80, 140, 144, 160, 224, 249, 256-59, 268-69, 306, 313-24, 326, 355, 373, 393 inter, 373, 395 macro, 249, 261, 269, 280 micro, 258-59, 269 subjectless experience, 222 substance, 26, 53, 55, 64, 67, 78-80, 83, 93, 164, 252, 358, 360, 373, 375-76 immaterial, 373, 375-76 supervenience, 52, 59, 61-62, 65-66, 115, 119, 132-34, 145, 227 strong, 61-62, 68-69, 134, 160 weak, 65, 68 Swinburne, Richard, 376 synoptic hierarchy. See synoptic pyramid synoptic pyramid, 9, 153-54, 166-69, 172, 175 system, computational, 55, 59, 168 Taliaferro, Charles, 14, 379, 382, 384-85 Thales, 1 theism, 14, 369, 370-72, 380, 384-85, 399 dualistic-panpsychistic, 384 theory of natural individuals, 9, 153-54, 161, 164, 167, 170 Tolstoy, Leo, 226 Tong, David, 222-23

Tononi, 67, 167, 171, 193, 201, 212 truth, necessary, 5 Tumulka, Roderick, 330 Turing machine, 58 twin earth, 288 Tye, Michael, 311, 372, 374, 376

ultimate-carrier problem, 56, 65 unity of consciousness, 144, 179, 259–61, 262, 279, 321 universalism. *See* composition: unrestricted phenomenal unpredictability of emergent properties, 341 structure, 13, 342

Van Cleve, James, 133, 310, 313, 321–23, 330 Van Gulick, Robert, 133, 331, 335 Van Inwagen, 12, 296, 301, 306 von Goethe, Johann Wolfgang, 215–16

Wager, Khai, 8 Weinberg, Steven, 86 what it's like, 58, 133, 137, 149, 153, 156, 169, 292, 307
Wheeler, John A., 101
White, Stephen, 230
Whitehead, Alfred North, 51, 66, 68–69, 80, 82, 94, 106, 166–67, 237–38
Williams, J. Robert, 301
Williford, Kenneth, 110
Wilson, Rob, 108, 336
Wimsatt, William, 227
Wishon, Donovan, 358, 366
Wittgenstein, Ludwig, 279, 379, 390, 403
Wong, Hong Yu, 15, 65–66, 68–69, 247
world-knot, 105–6
Yablo, Stephen, 45, 327–28

Zahavi, Dan, 110, 403 zombie, 22–23, 28–29, 36, 38, 43, 52, 57–58, 61, 135, 142–43, 146, 188, 226–27, 243, 269–70, 300, 319 awareness, 11, 269 o-, 45 t-, 45