Contemporary Debates in Philosophy of Mind
Contemporary Debates in Philosophy

In teaching and research, philosophy makes progress through argumentation and debate. *Contemporary Debates in Philosophy* presents a forum for students and their teachers to follow and participate in the debates that animate philosophy today in the western world. Each volume presents pairs of opposing viewpoints on contested themes and topics in the central subfields of philosophy. Each volume is edited and introduced by an expert in the field, and also includes an index, bibliography, and suggestions for further reading. The opposing essays, commissioned especially for the volumes in the series, are thorough but accessible presentations of opposing points of view.

1. Contemporary Debates in Philosophy of Religion *edited by Michael L. Peterson and Raymond J. VanArragon*
2. Contemporary Debates in Philosophy of Science *edited by Christopher Hitchcock*
3. Contemporary Debates in Epistemology *edited by Matthias Steup and Ernest Sosa*
4. Contemporary Debates in Applied Ethics *edited by Andrew I. Cohen and Christopher Heath Wellman*
5. Contemporary Debates in Aesthetics and the Philosophy of Art *edited by Matthew Kieran*
6. Contemporary Debates in Moral Theory *edited by James Dreier*
7. Contemporary Debates in Cognitive Science *edited by Robert Stainton*
8. Contemporary Debates in Philosophy of Mind *edited by Brian P. McLaughlin and Jonathan Cohen*
9. Contemporary Debates in Social Philosophy *edited by Laurence Thomas*
Contemporary Debates in Philosophy of Mind

Edited by

Brian P. McLaughlin and Jonathan Cohen
We dedicate this volume to Judy H. McLaughlin and Liza Perkins-Cohen.
# Contents

Notes on Contributors ix
Introduction Jonathan Cohen xii

## PART I MENTAL CONTENT

### Is There a Viable Notion of Narrow Mental Content? 3
1 Cognitive Content and Propositional Attitude Attributions Gabriel Segal 5
2 There Is No Viable Notion of Narrow Content Sarah Sawyer 20

### Is Externalism about Mental Content Compatible with Privileged Access? 35
3 Externalism and Privileged Access Are Consistent Anthony Brueckner 37
4 Externalism and Privileged Access Are Inconsistent Michael McKinsey 53

### Is the Intentional Essentially Normative? 67
5 Resisting Normativism in Psychology Georges Rey 69
6 Normativism Defended Ralph Wedgwood 85

### Is There Non-Conceptual Content? 103
7 The Revenge of the Given Jerry Fodor 105
8 Are There Different Kinds of Content? Richard G. Heck Jr 117

## PART II PHYSICALISM

### Is Non-Reductive Materialism Viable? 141
9 Everybody Has Got It: A Defense of Non-Reductive Materialism Louise Antony 143
10 The Evolving Fortunes of Eliminative Materialism Paul M. Churchland 160
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Should Physicalists Be A Priori Physicalists?</td>
<td>183</td>
</tr>
<tr>
<td>11 A Priori Physicalism</td>
<td>185</td>
</tr>
<tr>
<td>12 On the Limits of A Priori Physicalism</td>
<td>200</td>
</tr>
<tr>
<td>Is There an Unresolved Problem of Mental Causation?</td>
<td>225</td>
</tr>
<tr>
<td>13 Causation and Mental Causation</td>
<td>227</td>
</tr>
<tr>
<td>14 Mental Causation, or Something Near Enough</td>
<td>243</td>
</tr>
<tr>
<td>PART III THE PLACE OF CONSCIOUSNESS IN NATURE</td>
<td>265</td>
</tr>
<tr>
<td>Is Consciousness Ontologically Emergent from the Physical?</td>
<td>267</td>
</tr>
<tr>
<td>15 Dualist Emergentism</td>
<td>269</td>
</tr>
<tr>
<td>16 Against Ontologically Emergent Consciousness</td>
<td>287</td>
</tr>
<tr>
<td>Are Phenomenal Characters and Intentional Contents of Experiences Identical?</td>
<td>301</td>
</tr>
<tr>
<td>17 New Troubles for the Qualia Freak</td>
<td>303</td>
</tr>
<tr>
<td>18 A Case for Qualia</td>
<td>319</td>
</tr>
<tr>
<td>Is Awareness of Our Mental Acts a Kind of Perceptual Consciousness?</td>
<td>333</td>
</tr>
<tr>
<td>19 All Consciousness Is Perceptual</td>
<td>335</td>
</tr>
<tr>
<td>20 Mental Action and Self-Awareness (I)</td>
<td>358</td>
</tr>
<tr>
<td>Index</td>
<td>377</td>
</tr>
</tbody>
</table>
Notes on Contributors

**Louise Antony** is Professor of Philosophy at the University of Massachusetts, Amherst. In her research she attempts to develop naturalistic accounts of meaning, knowledge, and agency that square with our scientific understanding of the mind. She is the author of numerous articles in the philosophy of mind, epistemology, and feminist theory, and has co-edited two collections of original essays, *Chomsky and His Critics* (with Norbert Hornstein) and *A Mind of One’s Own: Feminist Essays on Reason and Objectivity* (with Charlotte Witt).

**David Braddon-Mitchell** is Reader in Philosophy at the University of Sydney. He has published papers in the philosophy of mind and metaphysics in *Noûs*, the *Journal of Philosophy, Mind*, the *Australasian Journal of Philosophy, Philosophical Studies*, the *British Journal for the Philosophy of Science, Erkenntnis, Synthese*, and various others. He is author, with Frank Jackson, of *The Philosophy of Mind and Cognition*.

**Anthony Brueckner** is Professor of Philosophy at the University of California, Santa Barbara. He has written articles about skepticism, transcendental arguments, content externalism, self-knowledge, theories of justification, anti-realism, personal identity, and the metaphysics of death.

**Paul M. Churchland** is Professor of Philosophy and holds the Valtz Chair at the University of California, San Diego. He is the author of *Scientific Realism and the Plasticity of Mind, Matter and Consciousness*, and *Neurophilosophy at Work* (forthcoming). His research lies at the intersection of cognitive neuroscience, epistemology, and the philosophy of mind.

**Jonathan Cohen** is Associate Professor of Philosophy at the University of California, San Diego. He works on topics in philosophy of mind, language, and perception, particularly as these are informed by the cognitive sciences. Many of his articles in recent years have concerned the metaphysics of color.
Jerry Fodor shares an appointment between the Philosophy Department at Rutgers and the Rutgers Center for Cognitive Science. He is currently at work on a book about the language of thought.

Richard G. Heck Jr is Professor of Philosophy at Brown University. He works primarily on philosophy of language, logic, and mathematics, and has written extensively about the work of Gottlob Frege.

Frank Jackson is Distinguished Professor of Philosophy at the Australian National University.

Jaegwon Kim is the William Perry Faunce Professor of Philosophy at Brown University. Among his recent publications are Physicalism, or Something Near Enough (2005) and Philosophy of Mind (2nd edn., 2006).

Barry Loewer is Professor at Rutgers, Director of the Rutgers Center for Philosophy and the Sciences, and fellow at the Collegium Budapest. He is the author of many articles in the philosophy of physics, the philosophy of mind, metaphysics, and logic. He is finishing a book on the metaphysics of laws, causation, and chance.

Michael McKinsey is Professor of Philosophy at Wayne State University. He is the author of many articles in the philosophy of language, the philosophy of mind, and ethics. His work has primarily concerned the semantics of natural language, especially the meaning and reference of proper names, indexicals, and natural kind terms, as well as the meaning and logical form of cognitive ascriptions.

Brian P. McLaughlin is Professor of Philosophy at Rutgers University. He is the author of numerous articles in the philosophy of mind.

Martine Nida-Rümelin is Professor at the Department of Philosophy, University of Fribourg, Switzerland. Her published work focuses on issues related to the debate about the ontological status of consciousness, especially on phenomenal consciousness (phenomenal concepts, anti-materialist arguments) and on transtemporal identity of conscious individuals (first person thought and identity across time). She is at present directing a research project on the philosophy of color vision. In her more recent research she has worked on the phenomenology of agency and the problem of free will.

Christopher Peacocke is Professor of Philosophy at Columbia University and a Fellow of the British Academy. He is currently working on a book on reference, understanding, and reasons.

Jesse Prinz is Associate Professor of Philosophy at the University of North Carolina, Chapel Hill. His books and articles concern various aspects of the mind, including consciousness, concepts, emotion, moral psychology, and cultural cognition.
Georges Rey is Professor of Philosophy at the University of Maryland at College Park. He works primarily in the philosophy of psychology, particularly the foundations of cognitive science, and has written extensively on the nature of concepts, images, qualia, and consciousness. He is the author of *Contemporary Philosophy of Mind* (Blackwell, 1997), and the editor of the cognitive science entries for the *Routledge Encyclopedia of Philosophy* and (with Barry Loewer) of *Meaning in Mind: Fodor and His Critics* (Blackwell, 1991). He is at present working on a book on the role of intentionality and intentional inexistents in early visual and linguistic processing.

Sarah Sawyer is Associate Professor of Philosophy at the University of Sussex. Her research is in the philosophy of mind, epistemology, and the philosophy of language.

Gabriel Segal is Professor of Philosophy at King’s College, London. He is author of *A Slim Book about Narrow Content* (MIT Press, 2000) and co-author, with Richard Larson, of *Knowledge of Meaning* (MIT Press, 1995).

Sydney Shoemaker is Susan Linn Sage Professor of Philosophy Emeritus at Cornell University. He is the author of *Self-Knowledge and Self-Identity, Identity, Cause and Mind*, and *The First-Person Perspective and Other Essays*.

Michael Tye is Professor of Philosophy at the University of Texas at Austin. He is the author of several books on consciousness, including *Ten Problems of Consciousness*.

Ralph Wedgwood is Lecturer in Philosophy at the University of Oxford and a Fellow of Merton College, Oxford. He has published several articles on metaethics, epistemology, and related areas in metaphysics and the philosophy of mind. In 2007 he will publish a book, *The Nature of Normativity*, that will give a unified presentation of some of his thinking about these subjects.
Introduction

Jonathan Cohen

Philosophy of mind today is a sprawling behemoth whose tentacles reach into virtually every area of philosophy, as well as many subjects outside of philosophy. Of course, none of us would have it any other way. Nonetheless, this state of affairs poses obvious organizational challenges for anthology editors. Brian McLaughlin and I have attempted to meet these challenges in the present volume by focusing on ten controversial and fundamental topics in philosophy of mind. “Controversial” is clear enough: we have chosen topics about which there is not a settled consensus among philosophers. By “fundamental” we don’t mean that the issues are easy or that the approaches taken toward them are introductory. Rather, we mean that (i) the resolution of these topics has implications for other issues inside and outside philosophy of mind, and (ii) past rounds of debate have revealed these topics as underlying broader disagreements. We asked leading philosophers of mind to defend one side or another on these topics. The result is what you now have in your hands.

In the remainder of this introduction I’ll say something by way of explanation of the topics covered and attempt to say how the topics relate to one another.

Content

A first cluster of topics concerns the nature of mental content. To say that mental states have content is to say that they can be about other things: for example, my current belief that there is a coffee cup on my desk is about the coffee cup and the desk. That mental states can be about things is a striking fact about them, and one that distinguishes them from most entities in the world (e.g., atoms, rocks, tables, numbers, properties). Moreover, insofar as things other than mental states (e.g., words, some paintings, scientific models) can have content, many philosophers have followed Grice (1957) in maintaining that they do so only by deriving their content from that
of the mental states of the makers or users of these other things; thus, while a painting might also be about the coffee cup, the Grice-inspired thought is that it has this content only by virtue of the content of the painter's intentions (e.g., her intention to produce a painting that is about that particular coffee cup), which are of course mental states. If this general picture is right, then mental content is more fundamental than other sorts of content. But what sort of a thing is mental content? And how is it constituted? What makes it the case, for example, that my current thought is about a coffee cup rather than a palm tree or nothing at all? These and related questions lie at the heart of the first cluster of topics in this volume.

Our first topic in this cluster is best appreciated against the backdrop of work starting in the mid-1970s (e.g., Putnam, 1975; Burge, 1979) arguing that the content of a thought is not wholly determined by the internal state of the thinker’s brain. On the contrary, these writers argued for what has come to be called content externalism—the view that what a thought is about is partially determined by factors outside the head of the thinker, such as the thinker’s physical and social environment. In Chapter 1, Gabriel Segal argues against content externalism. More specifically, he argues that what he calls “cognitive content” – the kind of content invoked in psychological explanations and propositional attitude ascriptions – is not fixed externalistically. His claim is that, even if externalists are right that the extensions of public language words (e.g., “water”) are determined by factors outside the thinker’s brain, nonetheless the cognitive content expressed by such terms is (i) idiosyncratic to individuals (or even time-slices of individuals), and (ii) determined by factors inside their heads. If so, then cognitive content is best understood as a kind of narrow or individualist (as opposed to externalist/anti-individualist) content. Sarah Sawyer argues against this approach in Chapter 2. She argues that if cognitive contents were to float free from the shared meanings and extensions of the public language words we use to attribute contents, as Segal holds, then it would be a rare miracle if any verbal attribution ever succeeded in capturing anyone’s cognitive contents. And this, she claims, would make a mystery of the utility and ubiquity of our practice of making verbal ascriptions of psychological contents to others. Ultimately, she contends, proponents of narrow content have failed to appreciate the significance, force, and scope of extant arguments for content externalism.

A second issue connected with content externalism comes up in Chapters 3 and 4, and concerns privileged access about the content of our mental states. It seems deeply plausible that our access to the content of at least some of our thoughts has some sort of epistemic privilege. For example, it seems deeply plausible that if I take myself to be thinking about water, it is truly water (not coffee, not a palm tree, and not some clear, tasteless liquid other than water) that is the subject of my thought. However, in recent years philosophers have argued that content externalism poses a serious threat to this plausible idea. The thought here is that if, as per externalism, the contents of my thoughts depend on factors outside my head (including contingent facts about the existence of particular elements of my physical and social environment), then I won’t know what those contents are whenever I am ignorant about the relevant external factors. In Chapter 4, Michael McKinsey argues that privileged access and content externalism are indeed incompatible, and that we should respond to the incompatibility by giving up the former. Anthony Brueckner holds, in Chapter 3, that
the alleged incompatibility is merely apparent. He argues that, although content externalism entails that the content of my thought depends on contingent facts about my environment, it does not entail that my knowing the content of my thought requires knowing contingent facts about my environment: consequently, Brueckner holds, it is consistent with content externalism that I can know the content of my thoughts without having knowledge of contingent facts about my environment. Their debate raises important issues about exactly how to understand the entailments content externalism has about thinkers’ environments, and about how we should individuate thoughts.

The volume also contains debates on two other foundational debates about content: one about the alleged normativity of content and one about how best to think about non-conceptual content.

The debate about the normativity of content is joined in Chapters 5 and 6 by Ralph Wedgwood and Georges Rey. The issue here is whether intentional (contentful) mental states, such as beliefs, desires, the acceptance of inferences, and so on, are constitutively tied to “normative” properties such as value, goodness, and, in particular, rationality. Such normative properties are traditionally contrasted against the “descriptive” properties one finds invoked in the natural sciences. Thus, this debate has important implications for the question of whether the standard explanatory apparatus of the natural sciences can provide a complete account of contentful mental states.

Wedgwood argues that the intentional is essentially normative. He holds that intentional states are constituted by concepts, and he argues that the best theory of concepts has them constitutively linked to the normative. In particular, Wedgwood is attracted by a two-factor theory of concepts according to which each concept is constituted by (i) its correctness condition together with (ii) “certain basic principles of rationality that specify certain ways of using the concept as rational (or specify certain other ways of using the concept as irrational)” (p. 86). Thus, for example, on this account, we might understand the concept of logical conjunction as constituted by (i) the systematic contribution made by AND to the truth conditions of the complex contents in which it appears (its correctness condition) together with (ii) a principle specifying that (inter alia) the inference from (P AND Q) to P is rational while the inference from P to (P AND Q) is not. Insofar as this conception of the constitution of concepts ineliminably invokes notions of rationality, it results in an essentially normative view of the intentional; but Wedgwood argues that his is the most plausible view of concepts, so we should embrace the latter result.

Rey argues against Wedgwood’s view in Chapter 5, and urges that our best scientific and philosophical accounts of mentality support a non-normative (“merely” descriptive) understanding of the intentional. Among the many complaints he levels against normative theories of the intentional, Rey worries (i) that there is no serious account of just which norms characterize particular concepts; (ii) that normative accounts of concepts don’t do justice to the portions of our mental lives that don’t seem to be governed by rational norms at all; and (iii) that even where applicable, such accounts give at best a superficial account of our mental lives. Rey suggests that Wedgwood and other proponents of an essentially normative account of the intentional base their view largely on intuitions about which rational inferences they are disposed to make involving particular concepts; but, while allowing that these intuitions are often
widely and deeply held, he echoes Quine (1951) in worrying that their wide and deep support may show only that these inferences are deeply ingrained (as opposed to concept-constitutive, as Wedgwood claims). If so, Rey points out, then such intuitions (despite being widely and deeply held) should not be taken as revealing the nature of our concepts; but if taking these intuitions to be concept-constitutive really is the source of the view that concepts are normative, then Rey’s worry threatens the case for the essentially normative character of the intentional.

In Chapters 7 and 8, Jerry Fodor and Richard Heck take on the topic of non-conceptual content. Discussion of this issue has centered in part on issues about perceptual justification. Many writers have thought that the best way to understand how perception justifies belief is by attributing content to perceptual states – thus, for example, my belief that there is a coffee cup on the desk would receive its justification from being appropriately related to a perceptual state with the very same content (that there is a coffee cup on the desk). But (a suitably generalized version of) this picture threatens to impose high cognitive demands on perception: it seems to require that our perceptual contents, in order to play any justificatory role, must be fully conceptualizable (see Sellars, 1956, for a famous articulation of this worry). But many philosophers have felt that this demand is unreasonable – for example, because it threatens the idea of a preconceptual “given” that could justify belief, or because it threatens to rob the possibility of perceptual justification from non-human animals and human infants.

Some philosophers of mind have maintained that the best response to these threats is to credit perceptual states with a special kind of “non-conceptual content” – content whose tokening is both (i) suited to justify the conceptual content of beliefs, and (ii) not dependent on sophisticated conceptual capacities of the perceiver. The problem for theorists sympathetic to this move is to provide an informative characterization of this hypothesized non-conceptual content, and then to give reasons for believing there is any mental content satisfying that characterization.

This is where both Fodor and Heck begin in their essays for the present volume. Both accept the existence of non-conceptual content (in this sense they are both giving kinds of “yes” answers to the question “is there non-conceptual content?”); but they differ in how they understand what it is, and how to distinguish non-conceptual content from conceptual content. In his contribution, Fodor spends most of his effort massaging the philosophical question “is there non-conceptual content?” into a form that makes it susceptible to answers by empirical psychology. In particular, Fodor holds that a mental state is conceptual if and only if it is an instance of representation-as, and he takes it that such states count as bearing content in virtue of the information they carry about the world. Thus, for Fodor, the existence of non-conceptual content hinges on the evidence in favor of mental states that are contentful (in the informational sense) but not instances of representation-as. But, Fodor argues, there is ample psychological evidence of states of this kind, so we have reason to accept the existence of non-conceptual content. Heck also spends much of his essay trying to get clear on what sort of a thing non-conceptual content might be. According to Heck, it is structural features of a contentful state that make it conceptual or non-conceptual: the state will count as conceptual if it has constituent structure, and non-conceptual if not. This criterion allows Heck (unlike Fodor) to accept that
instances of representation-as could be non-conceptual – namely, by lacking the right sort of constituent structure. Indeed, Heck argues that, on this way of making the distinction, the best accounts of perceptual content entail that it is non-conceptual.

**Physicalism**

A second, more ontological, cluster of topics taken up in this anthology concerns the relationship of mental states to the physical. Discussion of these topics is often organized around the physicalist/materialist hypothesis that everything (a fortiori, everything mental) is physical. With a few notable exceptions, contemporary philosophers of mind are generally sympathetic to some version of this hypothesis. But there is a startling lack of consensus about the details.

One way in which consensus is left behind is over the question of whether the best version of physicalism is reductive, non-reductive, or eliminativist – an issue taken up by Paul Churchland and Louise Antony in Chapters 9 and 10. Eliminative materialism, which is defended by Churchland in Chapter 10, is the view that our mental lives can be fully characterized by the (physical) kinds of neuroscience, and that putative psychological kinds such as belief, pain, and desire should be discarded as posits of a failed and outdated explanatory framework. Antony, in contrast, sees an important scientific role for such psychological kinds. Indeed, she wants to insist on a non-reductive materialism that preserves a place for these kinds without reducing them to (identifying them with) physical kinds. The best-worked-out version of non-reductive materialism, endorsed by Antony in Chapter 9, is the so-called functionalist view according to which mental types are understood in terms of their causal profiles; on this view, for example, a state might count as a pain if it is caused by damage to its host organism and causes “ouch”-utterances and avoidance behavior (as it might be), no matter what its physical realization. Antony argues that this kind of non-reductive materialism is not only viable, but preferable to reductive or eliminative materialism in that it better respects the reality and causal/explanatory centrality of psychological state types (Antony calls this “psychological realism”), and the distinctness of the phenomena and explanations of psychology from lower level (e.g., neuroscientific) phenomena and explanations (she calls this “the autonomy of psychology”). Churchland defends eliminative materialism in Chapter 10 by claiming that non-reductive materialism has been oversold. In particular, he urges that the most popular functionalist versions of the view have failed to meet the promises made on their behalf, while eliminativist materialism turns out to be more plausible than many have allowed.

A second way in which consensus about physicalism is left behind turns on a contrast between a priori and a posteriori versions of the thesis. It is widely accepted that if physicalism is true, then the physical determines the mental. In contrast, it is deeply controversial whether the determination of the mental by the physical is a priori or a posteriori. What is in dispute here is not the epistemic status of physicalism itself (all sides agree that if physicalism is true then it is posteriori), but instead the epistemic status of the determination of the mental by the physical that is implied by physicalism.
In Chapter 11, Frank Jackson defends a priori physicalism on epistemological and semantic grounds. In particular, he alleges that, if a posteriori physicalism were true, then this would undercut our warrant for adopting physicalism in the first place, and would also leave ordinary users of mental predicates without an understanding of what we are saying in ascribing such predicates (e.g., pains and beliefs) to each other. In Chapter 12, Brian McLaughlin first tries to cast doubt on the a priori physicalist thesis that the totality of truths of ultimate physics (in conjunction with a certain minimality thesis) will imply a priori all the truths couched in our everyday, vernacular physical vocabulary (e.g., the truth that water freezes at 32 °F). He maintains that it is an open question whether the concepts expressed by terms in our ordinary physical vocabulary will bear the kinds of a priori links to the concepts expressed in the vocabulary of ultimate physics that would be needed to underwrite such a priori implications. He then appeals to the conceivability of certain absent and inverted qualia cases – cases that have traditionally been used against physicalism itself – to argue that the links between our phenomenal concepts and physical/functional concepts are likewise non-a-priori. He holds that the phenomenal is indeed determined by the physical since phenomenal state types are identical with (broadly) physical/functional state types; but he contends that such type identity claims are warranted on a posteriori grounds of overall coherence and theoretical simplicity, and that they are not a priori implied by the totality of physical truths (and the minimality thesis).

One reason that the debate about a priori physicalism is so important is that it is intimately connected with the question of how far (certain kinds of) conceivability can be taken to reveal what is possible, which in turn bears directly on the modal commitments of physicalism. For we seem to be able to conceive of our world as one in which the physical facts fail to determine the mental facts. Now, if the determination required by physicalism comes with a priori knowability, then we should presumably be able to know a priori whether there is such necessary determination or not. But we have said that a priori reflection seems to leave open the possibility that there is no determination of the mental by the physical, which is to say that it tells against the claim that such determination is a priori. Thus, if we regard physicalism as requiring a priori determination, then what we can conceive poses a prima facie threat to its truth – a threat that has been regarded as fatal by at least some prominent philosophers of mind (e.g., Chalmers, 1996). On the other hand, if we regard the determination entailed by physicalism as a posteriori determination, we won’t regard the conceivability of differences in mental facts without differences in physical facts as a decisive objection to physicalism.

Another ontological dispute connected with physicalism concerns mental causation – causation by mental states. Ordinary action explanations (e.g., the explanation of why I drained the glass of water that cites my desire for liquid) bring out our pretheoretical commitment to the idea of causation by mental states. Unfortunately, it is unclear how to understand what this commitment amounts to. Part of the difficulty has its source in a more general controversy over the nature of causation (for example, between counterfactual, nomological, and productive approaches to causation). But there are difficulties particular to mental (or at least higher-level) causation as well. Perhaps the most widely discussed of these is the problem of explanatory exclusion,
pressed at the end of Chapter 13 by Jaegwon Kim. Kim worries that if every physical event has a sufficient physical cause, then there is no causal work left over for the mental to do. Kim takes this to show that either the mental is without causal efficacy (mental events would then be entirely epiphenomenal) or that the mental must be reductively identified with the physical. In Chapter 14, Barry Loewer disagrees with Kim’s assessment. He argues that we have the materials we need for understanding mental causation, unless we insist on a “productive” understanding of causation that he thinks is eschewed in science. Thus, he responds to Kim’s exclusion concerns by arguing that it is based on mistaken metaphysical presuppositions about causation. The upshot of Loewer’s chapter is that, while there may be unresolved problems about causation itself, there are no further outstanding problems about mental causation in particular.

The Place of Consciousness in Nature

The ontological debates surveyed so far are directed at issues about mental states, generally speaking. But in the last decade or so philosophy of mind has seen a renewed focus on ontological issues about consciousness in particular. Many philosophers have found consciousness to be especially resistant to explanation in physicalist terms, and this has raised profound concerns about its place in nature.

For example, some thinkers have thought that consciousness, unlike the rest of mentality, is ontologically emergent from the physical – that it is something fundamentally new and different from the physical. Thus, in Chapter 15, Martine Nida-Rümelin argues that at some point in the historical evolution of life, certain bits of matter got arranged in a way that marked a fundamental break with what had come before (viz., mere physical stuff): new individuals that are conscious came into being where none had been previously. Nida-Rümelin’s motivation for this view is a sense of puzzlement that she shares with many other philosophers, and that Levine (1984) famously dubbed “the explanatory gap”: it seems extremely hard to see how or why a certain complex physically organized system should enjoy any conscious phenomenology rather than none, or should enjoy the particular conscious phenomenology it does rather than some other. Some have argued that the existence of this gap reveals more about our kinds of minds and the concepts they deploy than it does about the relationship of consciousness to the physical; for these thinkers, the explanatory gap is not evidence of the ontological emergence or non-physical status of consciousness. Nida-Rümelin, however, is unimpressed by this treatment of the explanatory gap. She suggests, instead, that we should take the gap, and our natural “astonishment” about consciousness seriously – and that the best explanation of why we are astonished is that the relation between consciousness and the physical is, after all, just as deeply astonishing as emergentism says that it is.

David Braddon-Mitchell opposes this and other forms of emergentism about consciousness, in Chapter 16. As he sees it, the appeal of emergentism is the hope of securing what is attractive about both physicalism (its integration of consciousness with the physical) and dualism (its recognition of the distinctiveness of consciousness vis-à-vis the physical). Thus, the emergentist claims that consciousness is a novel,
hence genuinely emergent, feature of the world (this is the dualist ingredient) that emerges from a physical basis (this is the physicalist ingredient). However, Braddon-Mitchell argues, the emergentist’s two opposing poles of attraction ultimately make her position unstable. For if the emergentist insists on the dualist-inspired claim that consciousness is distinct from the physical, she thereby loses the ability to explain the causal relations between the base and what emerges, and consequently is stuck with an unattractive epiphenomenalism. On the other hand, if she emphasizes the connections between consciousness and the physical base from which it emerges sufficiently to avoid charges of epiphenomenalism, it will turn out that consciousness is straightforwardly physical. Thus, Braddon-Mitchell claims, there is no coherent way for emergentists to have their cake and eat it.

Questions about the place of consciousness in nature come up again in a related debate between Michael Tye and Sydney Shoemaker in Chapters 17–18. Tye and Shoemaker would agree that when you consciously see a ripe tomato, or taste a chocolate soufflé, your experience represents the world in some particular way. Moreover, picking up on the content-externalist themes discussed in connection with Chapters 1–4, both these authors would agree that the representational properties of your experience are determined at least partly by factors outside your head. What divides Tye and Shoemaker is the question whether there is a further aspect of your experience – its phenomenal character (or, as it is sometimes glossed, the what-it’s-like-to-have-it aspect) – that is distinct from its representational properties and is determined entirely by factors inside the head. Shoemaker argues that there is such a further, internalist, aspect of experiences, and concludes that the representational and phenomenal properties of experiences are distinct. Tye, in contrast, argues that the phenomenal character of an experience is identical to that experience’s (externally determined) representational content. More particularly, he argues against the view that phenomenal character is entirely determined by factors inside the head, and against the view that phenomenal character is nonrepresentational.

This debate will, of course, interest anyone who wonders what an adequate characterization of conscious experience will look like. Moreover, in asking how far philosophical ideas about content can be pressed in the service of explaining consciousness, it bears on the question whether we can reduce one philosophical problem to another. This last point is especially important because, as remarked above, many philosophers have been baffled by the problem of how to integrate consciousness into a physicalist ontology; and while there has by no means been a convergence on a single physicalist theory of content, it has seemed to many that the outstanding problems about content are (at least, by comparison to those about consciousness) solvable matters of detail.

A final debate about consciousness in this volume concerns conscious awareness of our own thought – the kind of awareness we have of what we are doing when we consciously deliberate, wonder, imagine, judge, and so on. In Chapter 20, Christopher Peacocke argues that we should conceive of our awareness of our own thought as a special form of action-awareness. Peacocke takes his inspiration from the (widely held) idea that subjects have a special, non-perceptual awareness of their own physical actions (say, the action of sitting, of kicking, etc.). Building on this idea, he maintains that subjects have a special, non-perceptual awareness of their own mental actions.
(say, the action of deliberating, of wondering, etc.), and takes this to motivate the view that awareness of thought is a species of action-awareness. Peacocke maintains that this conception of conscious thought not only provides the right way to think about the metaphysics, phenomenology, and epistemology of an important species of awareness, but also sheds light on related questions about self-knowledge and the first person. Jesse Prinz defends a sharply contrasting picture of conscious thought in Chapter 19 that gives a far more important role to perception. As his title suggests, Prinz holds the view that all consciousness, including consciousness of our mental acts, is perceptual consciousness. He defends this view by arguing that many of the putatively non-perceptual elements of our conscious mental lives are, on the best psychological and neuroscientific accounts, plausibly construed as perceptual after all. Moreover, since the existence of perceptual consciousness is accepted by all sides, he argues that parsimony should incline us against accepting a separate, non-perceptual form of consciousness to account for awareness of our own thoughts.

**Conclusion**

It would be impossible for an anthology like this one to touch on every topic, or even every important topic, or even every important and hotly disputed topic, in contemporary philosophy of mind. But the debates in this volume do, I think, give a fair sense of the current state of play with respect to many of the most fundamental and controversial issues in the subject. If they whet readers’ appetites for more of the subject, they will have served their purpose.

**References**


PART I
MENTAL CONTENT
IS THERE A VIABLE NOTION OF NARROW MENTAL CONTENT?
1 Background

Tyler Burge (Burge, 1979) has developed a very influential line of anti-individualistic thought. He argued that the cognitive content of a person’s concepts depends in part on their socio-linguistic environment. The argument of that paper centered around the example of a subject, now often called “Alf,” who does not know that by definition “arthritis” applies only to conditions of the joints. We are to imagine that Alf has many typical, mundane beliefs about arthritis: he believes that arthritis is a painful condition, that he himself has suffered from arthritis for years, that his arthritis in his wrist and fingers is worse than his arthritis in his ankles, and so on. One day he wakes up with a new pain and fears that his arthritis has spread to his thigh. On a twin Earth, the term “arthritis” is used more liberally than it is here, and applies to various rheumatoid conditions of bones and other sorts of tissue, and indeed is true of the condition that Alf has in his thigh. Twin Alf, being an inhabitant of this Twin Earth, has no concept of arthritis and so does not believe that his arthritis has spread to his thigh.

Burge would argue that (1) is true and (2) false:

1. Alf thinks that he has developed arthritis in his thigh
2. Twin Alf thinks that he has developed arthritis in his thigh

He would argue, more particularly, that (1) is true, and (2) false, on de dicto readings, where the content sentences specify the cognitive content of the attributed attitudes. Thus the cognitive contents of Alf’s and Twin Alf’s “arthritis” concepts differ. The difference is due to a difference between Earth and Twin-Earth experts’ opinion as to the meaning of the word-form “arthritis” in their respective languages. In particular, the opinions of Earth and Twin Earth experts fix different extension conditions for
the Earth and Twin Earth terms “arthritis.” This means that Alf’s “arthritis” concept is true of arthritis and arthritis only, while Twin Alf’s “arthritis” concept has a different extension condition.

And a difference of extension conditions entails a difference of cognitive content.

Burge’s articulation of the argument actually appears to involve commitment to a view that is significantly stronger than is required to establish the conclusion. The view is that, in typical cases, the words that a speaker uses are words of public language. A word of public language has a public meaning that is available to its different users. Speakers who are minimally competent with a word get to use the word with this public meaning. So, for example, if both Alf and a consultant rheumatologist say “Mr Jones has arthritis,” the words they utter express the same meaning – even though the expert believes that by definition “arthritis” applies only to swelling of the joints, while Alf does not believe this. Moreover, what is true of meaning is true of the content of propositional attitudes. When Alf and the doctor say “Mr Jones has arthritis” they express beliefs with the same cognitive content: what they both believe is that Mr Jones has arthritis.

Roughly following the terminology of Segal (2000), who was roughly following the terminology of Kaplan (1990) and Mercier (1994), we can label the two theses “weak consumerism” and “strong consumerism.” The idea behind “consumerism” is that a speaker is a consumer of public words with public contents: it is not for the speaker to produce his own meaning or content. I stipulatively use the terms as follows.

**Weak consumerism** is the (conjunctive) thesis that: (a) in typical cases, the extension conditions of the concept that a subject expresses by a term are partly determined by expert opinion, and (b) the cognitive content of a concept determines its extension conditions.

**Strong consumerism** is the thesis that: in typical cases, each term of public language has a unique cognitive content associated with it, and when a speaker uses the term, that is the cognitive content they express by it.

The two theses could come apart. It would be prima facie perfectly reasonable to endorse weak consumerism while rejecting strong consumerism. One might hold, for example, that Alf does not associate the very same cognitive content with “arthritis” as the expert, thus rejecting strong consumerism. But one might still hold that (1) is true and (2) is false on *de dicto* readings, because Alf has and Twin Alf lacks an “arthritis” concept that extends over arthritis and arthritis only.

The idea of adopting weak consumerism while rejecting strong consumerism receives some support from a natural response to the famous puzzle cases of Kripke (1979), of which the following is typical.

Ignacy Jan Paderewski (1860–1941, Polish) was both a great pianist and a renowned statesman. One can easily imagine a subject, call him “Barney,” who has heard of Paderewski the statesman and of Paderewski the musician, but who does not believe that they are one and the same person. Barney happens to believe that politicians typically lack musical talent. He is thus disposed sincerely to assent to both (3) and (4):
3. Paderewski had musical talent
4. Paderewski did not have musical talent.

Thus (5) and (6) both appear to be true:

5. Barney believes that Paderewski had musical talent
6. Barney believes that Paderewski did not have musical talent.

One might sensibly conclude that the classical, Fregean account of \textit{de dicto} propositional attitude attribution has to be given up. On the classical account, (5) and (6) entail that Barney believes contradictory Fregean thoughts. But, since Barney is being completely rational, this does not depict the situation correctly. One might, however, want to retain a fundamental aspect of Frege’s philosophy of mind. In particular, one might want to maintain that Barney’s dispositions in relation to (3) and (4) show that he associates different cognitive contents with his different uses of the name “Paderewski.” And that means that strong consumerism is false. There is at most one public-language term “Paderewski” (naming the musician and statesman), with at most one public content. But Barney uses different occurrences of the term to express different cognitive contents.

Now imagine a twin Earth with a twin Barney and a twin Paderewski. Twin Barney has beliefs about twin Paderewski that correspond to Barney’s belief about our Paderewski. But of course Twin Barney has no beliefs about Paderewski. And there is no interpretation under which (7) is true:

7. Both Barney and Twin Barney believe that Paderewski had musical talent.

And so, one might think, Barney’s and Twin Barney’s “Paderewski” beliefs do not have identical cognitive contents. So, one might conclude, extension is fixed by social facts, and cognitive content determines extension.

I will return later to the question of how we are to understand (5) and (6). I return now to Alf and arthritis. In Segal (2000) I articulated individualist responses to both strong and weak consumerism in relation to Alf. Allow me briefly to summarize.

The argument against strong consumerism, originally formulated in Loar (1987), works simply by turning the arthritis case into a Paderewski case. Suppose that Alf goes to France and there becomes competent with the term “arthrite.” He fails to realize that “arthrite” translates “arthritis.” He becomes disposed sincerely to assent to (8) and dissent from (9), a normal translation of (8) into French:

8. I have arthritis in my thigh

These linguistic dispositions show us that Alf has two different concepts with different cognitive contents that he expresses by the two terms “arthrite” and “arthritis.” But their public contents are the same. So, strong consumerism is false. The argument is defended in depth and detail in Segal (2000), and, in my opinion, can be made almost as irrefragable as an informal argument can get in philosophy.
My argument against weak consumerism (which is not almost as irrefragable as an informal argument can get in philosophy) went roughly as follows. When Alf says “arthritis” he does not mean arthritis. After all, that’s what he means when he says “arthrite.” Given this, there is no good reason to suppose that the concept he expresses by “arthritis” extends over arthritis and arthritis alone. And, given that, there is no compelling reason to suppose that Alf’s and Twin Alf’s concepts have different cognitive contents. Moreover, given that the concepts play identical roles in the two Alfs’ cognitive economies, there are good grounds for supposing that their concepts have the same cognitive content after all.

What I shall do now is consider a different, though closely related, argument of Burge’s that centers around a different but related type of example. The motivations for this consideration are: to show that the individualist responses to Alf are robust, to develop further the response to weak consumerism, and to explore some interesting issues that come to the fore in connection with the latter argument and example. After that, I will deal briefly with an objection to my approach.

2 Doubting Definitions

Burge (Burge, 1986) asks us to consider the case of a man, whom Burge calls “A” and whom I shall call “Arthur,” who has an iconoclastic theory about sofas. Arthur is competent in the use of the term “sofa,” having acquired it in a normal way. He is aware of what are considered to be standard truisms about sofas, such as that they are designed to be sat upon.

At some point, however, [Arthur] doubts the truisms and hypothesizes that sofas function not as furnishings to be sat upon, but as works of art or religious artifacts. He believes that the usual remarks about the function of sofas conceal, or represent a delusion about, an entirely different practice. [Arthur] admits that some sofas have been sat upon, but thinks that most sofas would collapse under any considerable weight and denies that sitting is what sofas are preeminently for. (p. 707)

We now imagine that a twin of Arthur’s, Twin Arthur, inhabits a twin Earth where the objects that he standardly calls “sofas” in fact are works of art or religious artifacts and are not made for sitting on. Most of them would collapse under any considerable weight. What Twin Arthur and his fellows call “sofas” are thus not sofas. Burge suggests that we call them “safos.” According to Burge, while Arthur has numerous propositional attitudes involving the concept of a sofa (Burge, 1986, p. 708) – such as that sofas are works of art or religious artifacts – Twin Arthur does not. Rather, Twin Arthur’s corresponding attitudes involve the concept of a safo. So, the contents of the subjects’ concepts depend in part on the natures of the artifact kinds in their local environment.

One obvious difference between this case and the “arthritis” case is that in this one it is not just expert opinion that varies across the twins’ environments, but also the nature of the artifact kinds involved. This will not be hugely important in what follows. But for bookkeeping purposes it requires another formulation of weak consumerism. I will call it “weak consumerismA,” “A” for “artifact”:
**Weak consumerism** is the (conjunctive) thesis that (a) in typical cases, the extension conditions of the concept that a subject expresses by an artifact term are partly determined by expert opinion and by the nature of the artifacts to which the term is actually applied; (b) the cognitive content of a concept determines its extension conditions.

A second major difference between the “sofa” and the “arthritis” case – a difference that Burge focuses on – is that Arthur is aware of the relevant definition, where Alf is not. While Alf just doesn’t know that, by definition, “arthritis” applies only to swellings of the joints, Arthur is aware that, at least according to some people, “sofa” is supposed to apply, by definition, only to objects made for sitting upon. But Arthur takes it that what are put forward as definitional remarks – such as “sofas are made for sitting on” – are incorrect. These remarks, he thinks, “conceal, or represent delusions about, a completely different practice.”

### 3 Paderewski Variations

I now deploy the Loarish argument against strong consumerism, in relation to Arthur. Let us indulge ourselves in another story. Arthur lives in London, England. He shares a flat with young man, Ban, who comes from a country in the far east called “Vantong.” Arthur and Ban become friends and spend much time in discussion of a multitude of topics. Arthur confides his views about sofas to Ban. Ban is at first skeptical. But he gradually becomes more open-minded in respect of Arthur’s iconoclastic views. Ban’s position is that he is quite certain that in his part of the world, sofas are really sofas. But he gradually becomes more and more sympathetic to Arthur’s views about sofas in the western world, until he is eventually convinced by them. However, Ban’s reasons for thinking that in Vantong sofas really are sofas are very convincing. His uncle is in the furniture business, he has visited a sofa factory, and so on. Arthur himself is at first skeptical about Ban’s claims. But after a while Ban convinces him. Thus they both arrive at the view that while the vast majority of sofas in the western world are religious artifacts or works of art, all the sofas in Vantong are genuine sofas.

Let us suppose that Arthur becomes reasonably competent in Vantong. The Vantong word for sofa, an exact translation of the English “sofa,” is “navid.” Arthur and Ban are agreed “navid” by definition applies only to items of furniture made for sitting upon and that anything that is not an item of furniture made for sitting on is automatically not a navid. They are both disposed sincerely to assent to the Vantong (10), which is an exact translation of (11) in normal English:

10 Navid yan-tse han saksak
11 All sofas are made for sitting on.

However they are of course happy to apply “sofa” to works of art or religious artifacts. And they are both disposed sincerely to dissent from (11).

If we were overhasty in our deployment of standard heuristics for attributing beliefs, we would conclude on the basis of their dispositions to assent to (10) that our
protagonists believe that all sofas are made for sitting on. And we would conclude from their disposition to dissent from (11) that they believe that it is not the case that all sofas are made for sitting on. We might then conclude that they are completely insane. A more charitable conclusion would be that they associate two different concepts, with two different cognitive contents, with the terms “sofa” and “navid.” Hence strong consumerism is false.

Moreover it looks as though, if there is such thing as the concept of a sofa, then it is the one they express by “navid” rather than the one they express by “sofa.” In that case, when Arthur says “sofa” he does not express the concept of a sofa. And if that is right, then it is not true that when Arthur uses “sofa,” he expresses propositional attitudes involving the concept of a sofa. Let us look into the matter in more detail.

4 De Re and de Dicto

What exactly does Arthur believe? Let us return to the time before he met Ban and focus on the attitude attribution (12):

Arthur believes that sofas are works of art or religious artifacts.

Two propositions are crucial to Burge’s argument for weak consumerism. The first is some utterances of (12) would be true on a de dicto reading, and the second is that, in such utterances, the word “sofa” extends over sofas and sofas alone. Further reflection casts doubt on these propositions. Utterances of (12) can be true on a de re reading. But a true de dicto reading is available only if “sofa” is being used in an idiosyncratic way. So let us further reflect.

When we are given the story of iconoclastic Arthur, we find it easy enough to accept (12). But that is on a de re reading like (13) or (14), which we also find unproblematic:

Arthur believes that long, upholstered seats are works of art or religious artifacts.

I think that there is a significant difference here between this case and the case of Alf. It is not natural to give (1) (“Alf thinks that he has developed arthritis in his thigh”) a de re interpretation. We are not thinking that Alf thinks of arthritis that he has developed it in his thigh. Part of the reason for this contrast is that it is very plausible that Arthur has various concepts of sofas that verify de re (de sofa) attributions. For example, he could think of sofas as instances of the same sort of artifact as that demonstrating a specific sofa. By contrast, Alf couldn’t think of arthritis as the same sort of disease as that, with that demonstrating the swelling of some joint. One reason for this is that “disease” doesn’t fix a principle of individuation that will collect the right extension: if the description fixes an extension at all, then it will be the disease that causes the arthritis – Paget’s disease, say (which can cause both arthritis and symptoms in the thigh) – rather than arthritis itself, which is a symptom of the disease. Relatedly, it would be hard to construct a plausible