

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

*Department of Philosophy, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0119, joncohen@aardvark.ucsd.edu

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. ??). 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 6, 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 (1953) 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. Fodor and Heck both accept the existence of non-conceptual content (this is why they are both listed as 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 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 *a 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 degrees Fahrenheit). 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 Braddon-Mitchell 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-a-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, he 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 souffle, 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

- Burge, T. (1979) "Individualism and the Mental," in P. French, T. Euhling, and H. Wettstein, eds., *Studies in Epistemology*, volume 4 of *Midwest Studies*

- in Philosophy*, University of Minnesota Press, Minneapolis.
- Chalmers, D. (1996) *The Conscious Mind: In Search of a Fundamental Theory*, Oxford University Press, New York.
- Grice, H. P. (1957) "Meaning," *The Philosophical Review* 66, 337–388.
- Levine, J. (1984) "Materialism and Qualia: The Explanatory Gap," *Pacific Philosophical Quarterly* 64, 354–361.
- Putnam, H. (1975) "The Meaning of 'Meaning' ," in K. Gunderson, ed., *Language, Mind, and Knowledge*, University of Minnesota Press, Minneapolis, 131–193.
- Quine, W. V. O. (1953) "Two Dogmas of Empiricism," in W. V. O. Quine, ed., *From a Logical Point of View*, Harper and Row, New York. Originally in *The Philosophical Review*, 1951.
- Sellars, W. (1956) "Empiricism and the Philosophy of Mind," in H. Feigl and M. Scriven, eds., *Minnesota Studies in the Philosophy of Science*, volume 1, University of Minnesota Press, Minneapolis, 253–329.