

David Chalmers, *The Character of Consciousness*. New York: Oxford University Press, 2010. Pp. xxvii, 596.

This book collects David Chalmers' work on consciousness from the publication of *The Conscious Mind* to the present. It is, for the most part, a collection of previously published essays — though there is substantial new material, both in the form of revisions and additions to the essays, and in the form of “afterwords” which are appended to five of the essays in the book.

It is well-known, but still worth saying, that Chalmers' work has defined the shape of contemporary work on consciousness. This is not just because Chalmers' views and arguments are widely influential — though of course they are. It is also because even philosophers of mind who disagree with Chalmers often define their own views in response to his and think about problems in the philosophy of mind in terms of the categories of views into which Chalmers has divided the space of possible solutions. It's extremely valuable to have Chalmers' work on these topics collected into a single volume.

The book is also testimony to the singular influence of Chalmers' work in another way, as it makes an impressive attempt to grapple with quite a lot of the vast literature which his work has inspired. This is sometimes daunting to the reader, as when a section begins with “We can start with fourteen putative counterexamples to CP” (which are followed by a bonus, fifteenth putative counterexample type). But it is also enormously useful to have the arguments in this sometimes confusing literature neatly taxonomized, and to have Chalmers' response to each clearly laid out.

Structurally, the book is a bit of a compromise. On the one hand, some attempt has been made to make this more than just a collection of independent essays; they are laid out in an order which traces a single coherent line of argument, and more effort than usual in a collection of essays has been made to draw explicit connections between the themes of the various chapters. On the other hand, there's more overlap between the chapters than one would expect from a standard monograph. This is a difficult compromise to carry off, but the book does it very well. The chapters are individually digestible, and the overlap doesn't detract too much from the experience of reading them seriatim.

...

The book starts off with a discussion of the nature of the problem of consciousness, distinguishing the “hard problem” — the task of giving an account of phenomenal consciousness — from other problems with which it can be confused.

From there, Chalmers moves on (in chs. 2-4) to discuss scientific attempts to grapple with this problem. Chalmers makes the case that a science of consciousness should not try to explain consciousness in terms of more fundamental things, but rather should treat consciousness as a fundamental element of reality alongside others. Chalmers suggests an analogy with the shift in 19th century physics to take electromagnetism as a fundamental aspect of reality alongside the ontology of previous physics. The analogy (as Chalmers is

aware) is not exact — the physical treatment of electromagnetism was motivated in part by purely extensional concerns of empirical adequacy, whereas non-reductive views of consciousness could be so motivated only if we were to include as empirical the results of journeys to zombie- and invert-worlds — but it is instructive as an example of a case in which taking a quality as fundamental rather than reducible is no bar to constructive science. In these chapters Chalmers also discusses a number of topics — like the concept of a neural correlate of consciousness — likely to figure in a science of consciousness, whether that science takes a reductionist or non-reductionist form.

Chs. 5-7 are in many ways the centerpiece of the book, and focus on the metaphysics of consciousness. Chalmers lays out his influential taxonomy of views in this area — types A, B, and C materialism, along with three anti-reductionist views: interactionist dualism (type D), epiphenomenalism (E), and monism (F) — and argues against the three types of materialist view. The core argument here is the conceivability argument. This argument has been the subject of a great deal of discussion; in a nutshell, and ignoring a raft of important distinctions, it goes like this: (i) it is conceivable that a world be a physical duplicate of our world but differ phenomenally from it; (ii) if this is conceivable, then it is possible; therefore (iii) it is possible. And if (iii) is true, materialism in all of its forms is false.

Much of this section is devoted to responses to objections to this argument, and I'll return to these below. But one might also wonder whether something like the conceivability argument can be used to argue against at least some non-reductive views. Chalmers discusses and effectively criticizes one way in which this argument might be used in a global argument against non-reductionism¹; but a more focused version of this strategy might be used to target monism.

According to monism (again, ignoring some complications), the phenomenal properties instantiated by subjects are to be explained by the intrinsic properties of the elementary particles of which they are composed. To see how the conceivability argument might be relevant to this view, consider a material conditional

$$I \rightarrow Q$$

where I is the collection of all the properties regarding the intrinsic natures of the particles composing some subject and Q the phenomenal property instantiated by the subject. Some such conditional must be necessary, or it follows immediately that monism is false. So is this necessary conditional a priori, or not?

If not, then it looks like we have an example of conceivability without possibility of a sort which would undercut the defense of premise (ii) of the sketch of the conceivability argument above.² So let's suppose that this conditional is a priori. But if it is a priori then one should be able to decide, a priori, between various hypotheses about one's I -

¹ See Chalmers' discussion of the "conceivability of materialism" argument at p. 180. Chalmers makes some points about monism relevant to what follows in pp. 133-7.

² I'm assuming that I -concepts — concepts of the relevant intrinsic particles — have the same primary and secondary intensions.

properties and P -properties on the basis of knowledge of one's own current phenomenal state. For suppose that n distinct I -properties, $I_1 - I_n$, necessitate, respectively, distinct phenomenal properties $Q_1 - Q_n$. Then for each such pair I^* , Q^* , the contrapositive of the relevant necessary conditional

$$\neg Q^* \rightarrow \neg I^*$$

will be a priori. Further, it is plausible that for any distinct phenomenal properties Q , Q^* , the biconditional

$$Q \text{ iff } \neg Q^*$$

will be a priori. It follows from this (plus the transitivity of a priori entailment) that given the information that I instantiate one of $I_1 - I_n$ and knowledge of my current phenomenal state, I'll be in a position to derive, a priori, the facts about the intrinsic properties. But it simply doesn't seem as though reflection on my own current phenomenal state puts me in a position, a priori, to decide between indefinitely many hypotheses about the intrinsic properties of the particles which compose me.³

One might object — mirroring the “missing concept” reply to the knowledge argument — that the reason why it seems that I'm in no position to carry out the relevant a priori inference is that I don't possess concepts (call them “ I -concepts”) for the intrinsic properties of, for example, quarks. Perhaps if I did possess the relevant concepts, then I could decide between the various hypotheses about my I -properties just on the basis of reflection on what it is like to be me right now.

But this sort of response seems unattractive, for a few reasons. Either possession of I -concepts requires a special sort of experience, or it doesn't. The first option seems implausible, because it's hard to see what sort of experience could be required. (This is a disanalogy with the missing-concept reply to the knowledge argument — there it is all too obvious what sort of experience is required.) But if we take the second option, then it should be possible for an ideal reasoner to form the relevant I -concepts only after reflection on his own phenomenal state — which leads to the result which we were trying to avoid. So it seems that I -concepts are impossible to possess and hence unthinkable — which seems a bit mysterious.

Second, this sort of response on the part of the monist sounds a bit like the following sort of speech, characteristic of type-C materialists:

Of course zombies are conceivable for *us*, since our understanding of the physical world is incomplete; but for an ideal reasoner in possession of the complete physics — and remember, such a person might grasp concepts for physical properties which we don't! — zombies would not be conceivable.

³ Though this may be less obvious if we think of the I -properties as themselves phenomenal properties rather than merely proto-phenomenal. Thanks to Leopold Stubenberg for helpful discussion here.

Our imaginary monist says:

Of course *we* can't know facts about a subject's *I*-properties a priori on the basis of facts about that subject's phenomenal properties; but an ideal reasoner — remembering of, course, that the ideal reasoner would possess *I*-concepts which we do not — would be able to know this.

To this sort of speech from the type-C materialist, Chalmers objects that physics tells us only about structure and dynamics, and that any future physics which did not lapse into monism would tell us only about structure and dynamics, and that we can know now that no truths about structure and dynamics will a priori entail the phenomenal facts (120-3).

But it's not obvious that an equally convincing objection can't be made to our monist's speech. After all, we already know (if the conceivability argument is any good) that phenomenal facts are not a priori entailed by facts about the intrinsic properties of *lots* of the things that compose me. For example, we know this about many intrinsic properties of organs, like the brain, and even about smaller things, like molecules — even if we don't know anything about the intrinsic properties of quarks, we do know some things about the intrinsic properties of molecules. Can't we just *see*, on the basis of consideration of examples of this sort, that facts about my phenomenal state will never be a priori deducible from facts about the intrinsic properties of things which compose me? (One might object: but the intrinsic properties of quarks might be *really different* than the known intrinsic properties of organs and molecules. But of course the facts about structure and dynamics revealed by future physics might also be *really different* than the corresponding claims made by present physics.)

This line of argument is only a problem for monism if Chalmers' elaboration of the conceivability argument is convincing. But if the foregoing is correct, dualists (whether epiphenomenalist or interactionist) who do find that argument convincing might deploy it against non-dualist non-reductive views as well as against various forms of materialism.

Let's return to the conceivability argument against materialism. A popular style of objection to this argument focuses on (ii), and points out that there are examples of truths which are necessary, even though conceivably false. Chalmers' initial response to this sort of move (type B materialism) has two prongs.⁴ First, we deploy the machinery of two-dimensional semantics, and argue that all cases where we have a (certain sort of ideal) conceivability without possibility are cases in which primary and secondary intensions of the relevant propositions diverge. (Roughly, the primary and secondary intension of an expression will come apart when there is at least one world *w* such that the reference of the expression with respect to *w* varies depending on whether we imagine ourselves learning that *w* is the actual world or whether we are thinking of *w* “counterfactually”, as a merely possible world.) Second, we argue that in the case of the conceivability/possibility claims used in the conceivability argument, we do *not* find a divergence between primary and secondary intension.

⁴ I say “initial” because Chalmers thinks that a suitably refined version of the conceivability argument does not depend on the assumption that the primary and secondary intensions of phenomenal concepts are identical; see, among other places, p. 153. I return to this point below.

Let's consider two different ways in which this argument might go wrong. First, it might be that our thoughts about phenomenal properties — our phenomenal concepts — do differ in their primary and secondary intension. Second, the relevant kinds of thought about phenomenal properties might essentially involve indexicality; in this case, the fact that we can't know a priori which phenomenal facts obtain on the basis of our knowledge of the physical facts would be a special case of the more general fact that we can't know indexical facts on the basis of knowledge of non-indexical physical facts.

So, it seems, the defender of the conceivability argument should want there to be an account of phenomenal concepts according to which at least some phenomenal concepts are (a) such as to allow no distinction between primary and secondary intension, and (b) non-indexical. And in Chs. 8-10, Chalmers distinguishes several different types of phenomenal concepts, and argues that one type — direct phenomenal concepts — has just the required features.

One worry here is that there appears to be some tension between (a) and (b), which can be brought out by considering Chalmers' claim that "[t]he lifetime of a direct phenomenal concept is limited to the lifetime of the experience which constitutes it" (272). This is a surprising claim about the possession conditions for any concept.⁵ Chalmers tries to allay worries here by pointing out that concepts can be possessed for very short periods of time, as when someone acquires a concept moments before death. While Chalmers is surely right about this, the observation doesn't seem on point — what seems odd about his claim is not that it entails that some phenomenal concepts are actually possessed for very short periods of time, but the modal claim that direct phenomenal concepts *can't* be possessed after the end of the relevant experience.

It is fair, I think, to expect some explanation of why direct phenomenal concepts can only be possessed while the phenomenal properties they are concepts of are instantiated. One very natural explanation would be that these phenomenal concepts are essentially indexical, and can for this reason only be possessed while the relevant phenomenal properties are around to be demonstrated. But this would violate constraint (b) on our account of phenomenal concepts.

Keeping (b) in place leaves us without an explanation for the fact that direct phenomenal concepts can only be possessed during the relevant experience, and therefore gives us some reason to abandon this claim.⁶ But this makes trouble for constraint (a). For suppose that I instantiate phenomenal property F , forming a direct phenomenal concept C_1 of F ; and then, a few moments later, instantiate F again, and this time form direct phenomenal concept C_2 of F . The question then arises: how are C_1 and C_2 related? Pretty clearly, they have the same secondary intension: after all, they are both direct

⁵ There are complications here about different uses of "concept" and "possessing a concept." Chalmers here is using "concept" as a term for mental representations rather than their contents, which is why I paraphrase his claim about the lifetime of direct phenomenal concepts in terms of a claim about the lifetime of the possession conditions for direct phenomenal concepts. See p. 254, note 3.

⁶ Though perhaps we could construct an explanation by developing Chalmers' idea that direct phenomenal concepts are "constituted by" the underlying phenomenal property.

phenomenal concepts of the same phenomenal property. But it seems as though their primary intensions differ. After all, the subject might legitimately wonder whether the phenomenal property represented by C_1 is the same as, or merely quite similar to, the phenomenal property represented by C_2 ; and it might really not be a priori for the subject that each presents the same phenomenal property.

But if the secondary intension of $C_1 =$ the secondary intension of C_2 , and the primary intension of $C_1 \neq$ the primary intension of C_2 , it follows that at least one of C_1 and C_2 , has a different primary and secondary intension — which violates constraint (a).

And, plausibly, once we see how this argument works, we can generate violations of (a) without the supposition that direct phenomenal concepts can be possessed after the end of the relevant experience — so long as it is possible for an ideal subject to simultaneously attend to two instantiations of the same phenomenal property (say, in distinct portions of his visual field), form direct phenomenal concepts of each, and, as above, be uncertain — and unable to resolve this uncertainty a priori — whether the relevant phenomenal properties are identical, or just quite similar.⁷

Chalmers might reply by pointing out that primary intensions are defined in terms of ideal a priori knowability, and say that, for an ideal subject, it will always be a priori (while they are instantiated) whether two simultaneously instantiated phenomenal properties are identical. But there are two worries about this move.

The first is just that it is not clear that this claim about ideal a priori knowability is, even if we limit ourselves to the synchronic case, correct; this can be brought out by imagining a phenomenal sorites-type case, in which the color presented in the left part of the subject's visual field is a constant shade of red while the color presented on the right is changing, ever so slowly, from orange, through the reds, to purple. Mightn't even an ideal reasoner be unable to figure out, a priori, when the phenomenal properties are the same?

The second is that this reliance on the gap between distinctness of primary intensions and facts about what seems cognitively significant to non-ideal subjects again raises worries about whether we should respect constraint (b). This is because Chalmers' central (though not only) argument for the claim that direct phenomenal concepts are non-indexical is an argument from the cognitive significance of identity claims involving one direct phenomenal concept and one explicitly demonstrative phenomenal concept. But if we are invited to doubt whether the cognitive significance (for us) of identity claims involving direct phenomenal concepts is a reliable guide to ideal a priori knowability, and hence a reliable guide to distinctness of primary intension, why should we not also doubt whether the cognitive significance of identity claims involving a direct phenomenal concept and an indexical phenomenal concept — each of which are, after all, concepts of the same phenomenal property which can't be possessed when that phenomenal property is not instantiated — is a reliable indicator of these concepts' having distinct primary intensions?

⁷ Chalmers discusses a related sort of “dual attention” scenario on p. 279.

These doubts are strengthened by suggestive parallels between the behavior of direct phenomenal and indexical concepts. It is characteristic of indexicals that they can be used to generate truths which are true whenever they are uttered or thought — like, for example, *I am here now* and *I am the thinker of this thought*. For any such truth, the conditional proposition that if it is thought, then it is true — for example,

If I am thinking that I am the thinker of this thought, then I am the thinker of this thought.

seems to be a priori. Further, since the antecedent of conditionals of this sort seems knowable on the basis of introspection alone, the consequent appears to be knowable on the basis of introspection + a priori reasoning. This is a point of contrast with identities which involve one indexical and one non-indexical term, like *I am Jeff Speaks*. By Chalmers' lights, (i) this proposition is not true whenever it is thought, (ii) the corresponding conditional, namely *If I am thinking that I am Jeff Speaks then I am Jeff Speaks*, is not a priori, and hence (iii) *I am Jeff Speaks* is not knowable on the basis of introspection + a priori reasoning.

With this contrast in mind, consider a proposition of the form *This phenomenal property is D*, where *D* is a direct phenomenal concept. If Chalmers is correct that direct phenomenal concepts are not indexicals, then we should expect this proposition to behave a like *I am Jeff Speaks*, and not like *I am the thinker of this thought*. But in fact, with respect to all of points (i)-(iii), *This phenomenal property is D* behaves more like the identity sentence containing two indexical singular terms. This proposition is true whenever it is thought, since, on Chalmers' view, one can only have thoughts involving a direct phenomenal concept while instantiating the phenomenal property it is a concept of, and demonstrative phenomenal concepts always pick out the phenomenal property one is instantiating.⁸ Furthermore, the fact that this proposition is true whenever thought is a priori — since the possession conditions for direct phenomenal and demonstrative concepts are knowable a priori — so the conditional

If I am thinking that this phenomenal property is D, then this phenomenal property is D.

will be a priori as well. Finally, since the antecedent of this conditional is knowable on the basis of introspection, the identity claim *This phenomenal property is D* is — like *I am the thinker of this thought*, and unlike *I am Jeff Speaks* — knowable on the basis of introspection + a priori reasoning.

These parallels between direct phenomenal and indexical concepts would vanish if we gave up the claim that direct phenomenal concepts can only be possessed while the

⁸ A complication: one can use the term “phenomenal property” either to pick out the total phenomenal state of a subject at a time, or to pick out the various aspects of that total phenomenal state. Here I'm using the term in the former way.

relevant phenomenal property is instantiated — but this, by the reasoning sketched above, seems to lead to the conclusion that there is a distinction between primary and secondary intension for direct phenomenal concepts (i.e., the falsity of constraint (a)).

Now, even if we end up taking this course, it is not obvious that this is anything other than an annoyance for proponents of the conceivability argument. As Chalmers points out (153), we can formulate a version of the conceivability argument which does not depend on the idea that direct phenomenal concepts are alike with respect to their primary and secondary intensions. But this version of the conceivability argument — like others — depends on the assumption (CP) that every conceivable proposition is verified by a canonical description of some centered metaphysically possible world. But the canonical descriptions of these centered possible worlds must exclude non-indexical expressions which differ with respect to their primary and secondary intensions.⁹ But if direct phenomenal concepts are among this excluded vocabulary, one might worry then that the canonical descriptions will be so sparse that there will be conceivable situations which are verified by no canonically described centered metaphysically possible world. And if this worry were realized, this might imperil the conceivability argument, since — given that, as Chalmers emphasizes, facts about which phenomenal properties are instantiated by subjects in a world are not a priori entailed by, for example, the physical properties of those subjects — zombie scenarios, and other relevant instances of CP, might be among them.

In Ch. 10, Chalmers presents an ingenious argument by dilemma against the view that the nature of phenomenal concepts can be used by the materialist to explain away the conceivability of these scenarios. Roughly, the dilemma is that either (1) it is conceivable that (functionally or physically identical) zombies not possess phenomenal concepts or (2) it isn't. If (1), then possession of phenomenal concepts raises just the same problems for the materialist as do phenomenal properties, so we can't solve problems posed by the latter by appeal to the former. If (2), then zombies must share our epistemic situation with respect to phenomenal properties — since our possession of phenomenal concepts was supposed to explain this epistemic situation. But zombies don't share our epistemic situation, since, not having the relevant phenomenal states, they, unlike us, cannot “conceive beings that lack phenomenal states that one actually has” (324).¹⁰

The argument that (1) is inconsistent with the view that conceivability intuitions can be explained away via phenomenal concepts is simple and devastating. But I don't see that the argument against someone who holds (2) is quite so convincing. One might hold

⁹ This is because the canonical vocabulary must include only semantically neutral expressions. See p. 551 and Chalmers (2006), §3.5.

¹⁰ Here as elsewhere, I'm focusing on the conceivability argument, and setting the knowledge and explanatory gap arguments against materialism to the side — if we focused on those arguments, different aspects of our epistemic situation would be relevant.

It's worth noting that this argument — unlike many other arguments in the book — would seem to require the conceivability of zombies, rather than just the conceivability of, for example, phenomenal inverters, since inverters could presumably conceive beings that lack phenomenal states that they have.

that our epistemic situation — in particular, the fact that we can conceive of beings like us physically but lacking our phenomenal properties — is jointly explained by the facts that (i) we possess phenomenal concepts and (ii) we instantiate the relevant phenomenal properties. It seems plausible that it is a priori that anyone who satisfies (i) and (ii) would share our epistemic situation. Zombies (on this line of response) satisfy (i) but not (ii), so it is no surprise that their epistemic situation differs from ours.

One might object that since the relevant features of our epistemic situation are a priori derivable from the conjunction of (i) and (ii) but not from (i) alone, the explanatory work is being done partly by phenomenal properties — and, since these phenomenal properties aren't a priori derivable from the physical properties, this might seem to push us back to horn (1) of the dilemma, since the explanatory gap between the physical and the phenomenal would seem to entail an explanatory gap between the physical and the conjunction of (i) and (ii).

But this is not obvious. If (i) and (ii) together a priori entail the facts about the conceivability of zombies, then (i) and (ii) together explain — in the relevant sense of “explanation” — the conceivability of zombies. Given that (as we're supposing) possession of phenomenal concepts is a priori entailed by the physical facts, then the conceivability of physical duplicates of ourselves without both properties (i) and (ii) does not seem to be a further fact which needs explanation — anything which explains the conceivability of zombies explains this. (This is by contrast with views which say that it is conceivable that physical duplicates of ourselves lack phenomenal concepts — the conceivability of such beings would be a further fact, and it would be hard to see how this fact could be explained by our possessing of phenomenal concepts.)

Chs. 11-13 turn away from the nexus of issues surrounding the metaphysics of consciousness and phenomenal concepts, and toward issues about the contents of experience. Chalmers defends an impure representationalist view, according to which experiences involve a certain propositional attitude relation to Fregean contents which involve not color properties themselves (as the Russellian would have it) but instead modes of presentation of those color properties, like *the property that normally causes phenomenally red experiences in me*. This sort of view is to be preferred over the popular combination of Russellianism with the view that colors are physical properties of surfaces for the familiar reason that the latter does not permit the possibility of spectrum inversion without misrepresentation.

At first, Chalmers' view can seem like a notational variant of a Shoemaker-style Russellian view which lets color experiences represent appearance properties, like the property of being disposed to cause phenomenally red experiences in me. Though the views are related, they are importantly different — not the least because, on Chalmers view, the complex italicized condition on reference above is not itself part of the content of experience, but merely singles out a primary intension — a function from worlds, considered as actual, to extensions — which is part of the content of experience. This is what enables Chalmers to say that his view, unlike Shoemaker's, avoids the conclusion that color experiences represent dispositional rather than intrinsic properties of surfaces.

This is part of the story about the content of experience; the other part is given in Ch. 12, where Chalmers argues that, in addition to these Fregean contents, experiences also represent primitive color properties which are (most likely) not actually instantiated. Chalmers suggests that both sorts of representational properties supervene on phenomenal properties, making them both species of “phenomenal content.” Unsurprisingly, given the foregoing, Chalmers does not think that this equivalence of representational and phenomenal properties will permit a reduction of phenomenal properties via a reduction of the representational properties. And, indeed, his Fregean view of content poses an extra dilemma for this sort of reductive program, since (as the sample mode of presentation given above illustrates) the relevant representational properties are specified partly in terms of phenomenal properties.

The last chapter of the book focuses on the unity of consciousness. Much of the chapter is given to characterizing the unity of consciousness. The claim that consciousness is necessarily unified in the relevant sense is then used to to construct an interesting argument against higher-order thought and representationalist-functionalist theories of consciousness, on the grounds that, were those theories true, it would be possible for a subject’s consciousness to fail to be unified.¹¹

Jeff Speaks
University of Notre Dame

References

Chalmers, David. 2006. “The Foundations of Two-Dimensional Semantics.” *Two-Dimensional Semantics*.

¹¹ Thanks for very helpful comments to David Chalmers, Leopold Stubenberg, and Casey O’Callaghan.