

Forthcoming in a special issue of the  
Journal of Consciousness Studies,  
ed. François Kammerer

## **CHALMERS'S META-PROBLEM**

David Rosenthal  
Philosophy and Cognitive Science  
Graduate Center, City University of New York

### **Penultimate Version**

#### **Abstract**

There is strong reason to doubt that the intuitions Chalmers's meta-problem focuses on are widespread or independent of proto-theoretical prompting. So it's unlikely that they result from factors connected to the nature of consciousness. In any case, it's only the accuracy of the problem intuitions that matters for evaluating theories of consciousness or revealing the nature of consciousness, not an explanation of how they arise. Unless we determine that they're accurate about consciousness, we mustn't assume that realism about consciousness incorporates them, or that denying them means denying that consciousness exists. And we must assess whether they're accurate independently of how they arise.

#### **I. Problem Intuitions**

The hard problem is "why and how do physical processes in the brain give rise to conscious experience." The meta-problem is "why we think consciousness poses a hard problem," that is, "why we think consciousness is hard to explain" (Chalmers 2018, 6; except as otherwise indicated references are to this article). Consciousness in this connection consists of conscious phenomenal states, that is, "states individuated by what it is like to be in them" (n. 2).

Chalmers sees a close connection between a solution to the meta-problem and the nature of consciousness. We

can reasonably hope that ... an explanation of our judgments about consciousness ... will give us insight into consciousness itself. Presumably there is at least a very close tie between consciousness and the mechanisms that generate reports about it (9-10).

The reports Chalmers has in mind here are those "that reflect our sense that consciousness poses a special problem" (11), what he calls "problem reports" (10). These problem reports contrast with "mundane reports such as 'I am feeling pain now'"

(12). Examples Chalmers offers of problem reports include, "Consciousness is hard to explain" (11), "I can't see how consciousness could be physical" (12), and "Explaining behaviour does not explain consciousness" (7). Those bear directly on explanation; others concern undetectable inversion of mental qualities and Mary's learning something new. Problem intuitions, then, are "dispositions to make specific problem reports and judgments" (11).

Chalmers's "mundane reports" are genuine reports; in saying that one is in pain one reports what one takes one's mental condition to be. But it's stretching things to see "Consciousness is hard to explain" as any kind of report, to say nothing of speculations about undetectable inversion or Mary. People do not say such proto-theoretical things spontaneously, as they report being in pain or seeing or hearing something.

Describing such remarks as reports suggests that they function as mundane reports do, to describe spontaneously some one's mental condition. But they don't. Reporting that one is in pain or that one perceives something typically results, in a way that's subjectively unmediated, from one's feeling pain or consciously perceiving something. Saying that consciousness is hard to explain, by contrast, or that your green could be my red requires some conversational setup, and likely some reflection. It is not a spontaneous report of one's state of mind.

Chalmers approvingly ascribes to Alvin Goldman (2000) an objection to higher-order-thought (HOT) theory that illustrates the difficulty. On that objection, "it is not clear how mere higher-order thoughts explain why we report mental states as being conscious" (38; emphasis here and throughout is Chalmers's). But we don't. Apart from an academic or other special context, people do not describe their conscious states as being conscious. Unless a contrast between conscious and unconscious mental states is under discussion, there's no point in describing a conscious state as being conscious.

Indeed, a spontaneous report that one is in pain or perceives something is by itself enough to understand that a state is conscious, without anybody's explicitly saying it is. For creatures with the relevant linguistic ability, conscious states are those one can report, seemingly independently of inference. And HOT theory explains this; reporting that one is in a mental state expresses a HOT that one is in that state. And the state is conscious because one is aware of being in the state by having that HOT.

These considerations cast doubt on Chalmers's suggestion that there is "a very close tie between consciousness and the mechanisms that generate [problem] reports about it." There likely are dedicated mechanisms that generate the mundane reports; indeed, verbally expressing HOTs would do just that (Rosenthal 2005). But the problem remarks are not straightforward reports; they result at least in part from background setup, prompting, and tacit proto-theoretic assumptions. So those factors will figure in generating the problem remarks, and such factors are no part of, and may well not reflect, the nature of consciousness.

Chalmers considers the charge that his so-called problem reports simply reflect “the intuitions of philosophers, and of a subclass of philosophers at that.” And he replies that “the central intuitions are widely shared well beyond philosophy,” and “that versions of many of these intuitions can be teased out of ordinary subjects” (13). “It is easy,” he urges, “to get ordinary people to express puzzlement about how consciousness could be explained in terms of brain processes” (7).

But even if true, this wouldn't show that the statements Chalmers has in mind are anything like spontaneous reports. Teasing puzzlement out of somebody is different from the person's reporting a subjectively unmediated take about a state of mind. And teasing out intuitions can be a double-edged sword. One can often elicit intuitions of one sort by describing things in a way that suggests those intuitions, but intuitions of a contrary sort by describing things differently. What intuitions get teased out typically depends on who is doing the teasing, and how.

In urging that his problem intuitions are natural and untutored, Chalmers also appeals to “a significant body of psychological data on the 'intuitive dualist' judgments of both children and adults” (7). But intuitive dualist judgments that the mental is not physical might mean not that the mental is nonphysical, but only that the mental is not simply physical; it might be a very special case of the physical (Rosenthal 1976). We sometimes distinguish As from Bs when the As are just a highly salient subclass of Bs that's of great interest, as when we distinguish humans from animals.

I know of no psychological data generated in a way that controls for this. And if the psychological findings are compatible with people seeing the mental as simply a strikingly special case of the physical, they have no bearing on Chalmers's claim that the problem intuitions are “widely shared well beyond philosophy.” This applies also to the work in experimental philosophy cited in note 7.

Even if there were some genuinely dualist tendency among those of the folk innocent of academic debate, it is notorious that what seems intuitively puzzling can change as scientific findings become second nature for us. It's difficult today to evoke a worry about Eddington's two tables—the alleged issue about how a macroscopic table can be composed of subatomic particles—or about vitalism. The physics and biology have become too deeply engrained in how we think about things for such worries to get traction. Why think the same wouldn't happen with consciousness once it becomes well understood how brain processes subserve consciousness and early education and surrounding culture make that second nature for us?

Chalmers occasionally appeals to concepts rather than intuitions; “our ordinary concept of phenomenal consciousness is not a functional concept” (50). But just as it's unclear that Chalmers's problem intuitions are widely shared and arise without proto-theoretic prompting, it's unlikely that those unexposed to such prompting have any concept of phenomenal consciousness at all. Perhaps Chalmers means some more commonsense concept, but it's unclear what that might be.

The meta-problem is how to explain why we have problem intuitions. Since it's questionable how widespread Chalmers's problem intuitions are and whether they mainly occur only when elicited by describing things in special ways, it's also questionable whether there is a meta-problem. If the problem intuitions are teased out when one is exposed to a particular way of describing things but don't typically occur without that, they aren't sufficiently widespread to call for explanation, and in any case not an explanation that might shed light on the nature of consciousness.

So before we can be confident that there is a meta-problem, we need far firmer reason to think that Chalmers's problem intuitions reflect the natural, spontaneous way everybody thinks about consciousness. Many in philosophy, psychology, and neuroscience, as well as many with no exposure to relevant academic work, disavow those intuitions. With their status in doubt, more is needed before we accept that there is a meta-problem, or perhaps even a hard problem.

## II. Evaluating Theories

Because Chalmers takes the problem intuitions as given, he sees realism about consciousness as incorporating the problem intuitions. "For a realist, judgments about consciousness systematically reflect the character of consciousness. ... [A] realist should expect that our judgments about consciousness are the way they are because consciousness is the way it is, or at least because the basis of consciousness is the way it is" (36). That would in turn sustain the close tie Chalmers sees between the hard problem and the meta-problem: "Whatever explains consciousness (the hard problem) should also play a central role in explaining our judgments about consciousness (the meta-problem)" (36).

Chalmers describes illusionism (Frankish, 2016) as the view that "consciousness is or involves a sort of introspective illusion." So illusionism seeks to explain the problem intuitions "as resting on an illusion" (8). Illusionism so understood resembles realism in taking the problem intuitions as given. Illusionism simply denies they're true, whereas Chalmers's version of realism takes them to be true.

So it's unsurprising that Chalmers tends to see realism and illusionism, so understood, as the only tenable views. Given his understanding of realism about consciousness, he espouses realism. Still, he writes, "I have more sympathy with [illusionism] than with most materialist views" (9). That's because illusionism, unlike other materialist views, takes the problem intuitions as given. Illusionism and realism, understood Chalmers's way, both hold that the problem intuitions are spontaneous and widespread, and that they're revealing about what consciousness is if it exists,

We've seen reason to doubt that the problem intuitions are widespread or occur without special prompting. But independently of that, there is a question about whether the problem intuitions say anything true about what consciousness is. And that's what

actually matters. If the problem intuitions aren't accurate about consciousness, realism about consciousness won't incorporate them, and denying that they apply to anything won't imply that consciousness doesn't exist. One can then be fully realist about consciousness and nonetheless deny that the intuitions say anything true about consciousness, as many who call themselves realists do.

Relying on the problem intuitions' being widespread and spontaneous and a strong tie between meta-problem and hard problem, Chalmers formulates a "meta-problem challenge for theories of consciousness. If a theory says that mechanism M is the basis of consciousness, then it needs to explain how mechanism M plays a central role in bringing about judgments about consciousness" (36). Any theory that doesn't endorse the problem intuitions gets disposed of by this challenge.

So, if integrated information is the basis of consciousness (Tononi 2008), it "should play a central role in explaining our judgments about consciousness" (36). But how could it? Similarly for global-workspace theory (Dehaene & Naccache 2001), Ned Block's biological theory (2009), and HOT theory (Rosenthal 2005). Other theories could readily be added. On what Chalmers calls a "principle" (36), we get to rule out any theory that doesn't explain the problem intuitions Chalmers thinks we have.

If special prompting and tacit proto-theoretical assumptions figure in the generation of the problem intuitions, those factors may well explain why some people have the intuitions. The meta-problem challenge would then be inapposite; we would explain the intuitions independently of explaining consciousness. The meta-problem won't constrain theories of consciousness unless it's the job of such a theory to explain the intuitions, and that's questionable if they might be explained by some other theory.

If the problem intuitions are accurate, perhaps the best explanation of their accuracy would be that they result from something in the nature of consciousness. And then perhaps a theory that explains consciousness should also explain the intuitions. But both those steps are highly speculative. Even if the intuitions arose because of something about consciousness, a satisfactory explanation of consciousness might not address that connection. And the intuitions might be accurate without any special or transparent connection to the nature of consciousness. Also, if the intuitions are accurate about consciousness, there is a hard problem, and any theory ruled out by the meta-problem challenge would likely also be unable to solve or dispel the hard problem. So the meta-problem would not constitute an additional hurdle.

And if there is a hard problem, addressing it is what's basic. So what matters is whether the problem intuitions are accurate about consciousness. Their being widespread and spontaneous might lend credence to their being accurate, but by itself cannot establish that. Their being accurate is independent of whether and why people think they are. So the benefit of casting things in terms of a meta-problem can only be that it takes the problem intuitions for granted, insulating them from challenge.

Sometimes a solution to a problem would, if accepted, result in things no longer seeming problematic. And it's sometimes urged that we should reject such a solution. Chalmers approvingly cites an example of this; Thomas Nagel rejects a functionalist solution to the problem of other minds in part because it "leaves it a complete mystery why [that problem] has ever bothered anyone" (1970, 403). The meta-problem challenge looks to be another example; we should dismiss any proposed explanation of consciousness that fails to preserve the problem intuitions.

This cannot be an appropriate constraint. Our being puzzled about something typically stems from our not seeing how to explain it; so seeing an explanation typically dispels puzzlement. Whatever the merits of a functionalist solution to the other-minds problem, it is no shortcoming that it results in the issue's no longer seeming problematic.

When an explanation works, we can recapture our puzzlement only by thinking about things as we did before accepting the explanation. As Nelson Goodman notes, when we attempt "to make the obscure obvious ... the reward of success is banality. An answer, once found, is dull" (1966, xix). Successful explanations make things fit together in a way that dissipates the sense of puzzlement. Rejecting an explanation simply because it has that effect precludes any explanation that could succeed.

### III. The Hard Problem

Casting thing in terms of a meta-problem suggests that we not only find consciousness problematic, but that this will persist come what may. And since Chalmers's version of realism incorporates the problem intuitions, he thinks it will persist. That's why we must evaluate whether those intuitions are widespread outside of special contexts, how much proto-theoretic prompting is needed to evoke them, and whether contrary intuitions that there's nothing especially problematic about consciousness can also be readily evoked with alternative prompting.

But even if the intuitions are widespread and spontaneous, they might not be accurate about the nature of consciousness. Chalmers's hypothesis is that they're widespread and spontaneous because some mechanism tied to consciousness generates them. And the operation of such a mechanism suggests to Chalmers that the intuitions are accurate. But no such mechanism could guarantee that the intuitions are accurate. Why we have the intuitions if we do is independent of whether they're accurate, and what matters is whether they're accurate.

Without widespread, untutored, spontaneous problem intuitions there is no meta-problem. But the hard problem, as Chalmers sometimes states it, is independent of our having problem intuitions. "[T]he core of the hard problem is posed ... by our experience of [mental] qualities: roughly, the distinctive phenomenal way in which we represent the qualities or are conscious of them" (30). And however one construes that, it's the nature of consciousness, not problem intuitions.

Chalmers might insist that the experience of the qualities somehow implies or otherwise involves the problem intuitions. But that's an extra step, not mentioned in the foregoing formulation of the hard problem. And if the intuitions aren't widespread or spontaneous outside special contexts, it's not likely that the experience of mental qualities does by itself imply or involve problem intuitions.

It's in any case unclear what that statement of the hard problem amounts to. It plainly relies on first-person access to qualitative mental properties, access Chalmers sometimes describes as "immediate acquaintance with consciousness" (46). And such access to mental qualities is certainly special. But being special does not imply being problematic. We need a clearer, more explicit description of what consciousness is and just what about it is problematic. We cannot rest with evocative allusions.

It may be tempting to meet this demand by urging that one simply look inside and reflect on whether what one introspectively finds there could be explained by appeal to brain functioning. Thus Chalmers writes that the question, "why are physical pain processes accompanied by the feeling of pain ... is as central a version of the hard problem as any," so that the problem can be posed "using the introspectively obvious datum alone" (54). And: "To generate the hard problem of consciousness, all we need is the basic fact that there is something it is like to be us" (49). These formulations are again independent of any specific problem intuitions.

But even if one does come to be puzzled by the "introspectively obvious datum," that does nothing to clarify what the datum is or what is puzzling about it. The foregoing formulations suggest that we can only allude to or gesture at it, and that the datum resists any straightforward description. This echoes what Block (1978, 281) answering the question what qualitative mental states with Louis Armstrong's unhelpful remark about what jazz is: "If you got to ask, you ain't never gonna get to know."

Having no informative description matters. The hard problem is "why and how do physical processes in the brain give rise to conscious experience." Without a clear, tolerably accurate description of what conscious experience is, we cannot begin to address that question or even evaluate whether doing so would be difficult. It's no help to be told that we know from looking inside what conscious experience is. We know in that way examples of conscious experience, not what it is.

Indeed, whenever we have no informative description of a phenomenon, we will be at a loss about how to explain it. This is entirely general; consciousness is not special in this way. The issue is not whether we can explain consciousness by appeal to brain processes; it is whether we can explain it at all without some reasonably informative description of what it is. How could an explanation take hold if we can't even describe the target phenomenon?

If we knew nothing of water but how it appears perceptually, it would seem puzzling how to explain it in terms of chemistry, and unlikely that we could. Similarly,

relying exclusively on our looking introspectively inside will lead to feeling at a loss about how to explain conscious experience by anything physical, or indeed by appeal to anything at all. But that doesn't show that anything is problematic; it shows only that we need an informative description of the phenomenon we want to explain.

Many questions arise in attempting to provide a reasonably clear and accurate description. Chalmers says we have "immediate acquaintance with consciousness." Is that acquaintance actually unmediated? Or is it just that we have a subjective sense that there's no mediation? After all, Chalmers speaks elsewhere only of "a sense of immediate knowledge" (n. 16). And what does that acquaintance consist in? Might it be simply an awareness of oneself as being in a relevant mental state, as HOT theory proposes? If not, why not?

These and other questions underscore the need for a clear, informative description of the phenomenon we're talking about. We cannot settle for a suggestive appeal to there being something it's like or alleged phenomenal intuitions intended to evoke a sense of puzzlement.

Having no informative description of some phenomenon, leads to a sense that it cannot be explained; this is not special to consciousness. Is that a solution to Chalmers's meta-problem? Can we explain whatever problem intuitions some have by their operating without an informative description of what they call consciousness? I won't venture an answer, since independently of that there is good reason to question the status of those problem intuitions.

#### **IV. Intuitions and Debunking**

Debunking arguments go from a claim that we have particular beliefs because of factors irrelevant to whether they're true to the conclusion that those beliefs are not justified. Such arguments are questionable. Even if we have particular beliefs because of factors irrelevant to their truth, other considerations could justify the beliefs. And justification aside, what matters is whether the beliefs are true. Debunking arguments get at truth in a highly indirect way if at all.

Focusing on a meta-problem, and hence on explaining why we may have particular intuitions about consciousness, is unsettling in a somewhat similar way. Even if the intuitions were widespread, spontaneous expressions of what we take the nature of consciousness to be, what matters is whether the intuitions are true, not why we have them.

Chalmers proposes that the intuitions are accurate because the best explanation of their being widespread is that they're generated by factors connected to consciousness. But their being so generated couldn't show that they're accurate. Even if the intuitions were widespread, independent of question-begging prompting, and due to factors connected with consciousness, they still might not be accurate.

The situation is parallel with debunking arguments; beliefs generated by a set of factors irrelevant to their truth need not be false. The origin of beliefs is independent of whether they're true. Chalmers anticipates that the best explanation of our having problem intuitions will involve a close tie to the nature of consciousness, and so provide good reason to think the intuitions are accurate. But the parallel with debunking arguments should give us pause. Until we have an actual explanation, the hypothesis of a close tie with the nature of consciousness is premature. Explaining why we have the intuitions, if indeed they were widespread and spontaneous, may tell us little if anything about the nature of consciousness.

Sometimes stepping back from a problem helps to make progress. That's the motivation for debunking arguments, and a focus on the meta-problem is similarly a kind of stepping back from the hard problem. But the problem intuitions that the meta-problem takes for granted concern the same issues that lead some to think there's a hard problem in the first place. Since settling whether the problem intuitions are accurate is the same as settling whether there's a hard problem, it's unclear that this case of stepping back can help.

Pretheoretic intuitions sometimes provide a useful guide for thinking about and investigating a phenomenon. But the last word must always include a theoretical account that does reasonable justice to common sense, revising it as needed, and that also explains how the target phenomenon fits with other relevant things we know. Intuitions without theory are blind.

## REFERENCES

- Block, Ned (2009), "Comparing the Major Theories of Consciousness," in The Cognitive Neurosciences, IV, ed. Michael Gazzaniga, Cambridge, Massachusetts, MIT Press, pp. 1111-1122.
- Block, Ned (1978), "Troubles with Functionalism," in Minnesota Studies in the Philosophy of Science, IX, ed. C. Wade Savage, Minneapolis: University of Minnesota Press, pp. 261-325.
- Chalmers, David J. (1995), "Facing up to the Problem of Consciousness," Journal of Consciousness Studies, 2, 3 (March): 200-219.
- Chalmers, David J. (2018), "The Meta-Problem of Consciousness," Journal of Consciousness Studies, 25, 9-10 (September-October): 6-61.
- Dehaene, Stanislas, and Lionel Naccache (2001), "Towards a Cognitive Neuroscience of Consciousness: Basic Evidence and a Workspace Framework," Cognition, 79, 1-2 (April): 1-37.
- Frankish, Keith (2016), "Illusionism as a Theory of Consciousness," Journal of Consciousness Studies, 23, 11-12: 11-39.
- Goldman, Alvin I. (2000), "Can Science Know When You're Conscious? Epistemological Foundations of Consciousness Research," Journal of Consciousness Studies, 7, 5 (May): 3-22.
- Goodman, Nelson (1966), The Structure of Appearance, 2nd edition, Indianapolis: Bobbs-Merrill.
- Nagel, Thomas (1970), "Armstrong on the Mind," The Philosophical Review, 79, 3 (July): 394-403.
- Rosenthal, David (2005), Consciousness and Mind, Oxford: Clarendon Press.
- Rosenthal, David (1976), "Mentality and Neutrality," The Journal of Philosophy, 73, 13 (July 15): 386-415.
- Tononi, Giulio (2008), "Consciousness as Integrated Information: a Provisional Manifesto," The Biological Bulletin, 215, 3 (December): 216-242.