

**The Hard Problem Isn't Getting Any Easier:  
Thoughts on Chalmers' "Meta-problem"**

Ben White

2021. *Philosophia*. 49: 495–506.

<https://doi.org/10.1007/s11406-020-00210-9>

**Abstract:** Chalmers' meta-problem of consciousness is the problem of explaining "problem reports"; i.e. reports to the effect that phenomenal consciousness has the various features that give rise to the hard problem. Chalmers (2018, 8) suggests that solving the meta-problem will likely "shed significant light on the hard problem." Against this, I argue that work on the meta-problem will likely fail to make the hard problem any easier. For each of the main stances on the hard problem can provide an account of problem reports, and we have no way of deciding which of these accounts gives the *correct* explanation of an individual's problem reports without presupposing a stance on the hard problem. We thus cannot determine which of the available solutions to the meta-problem is correct without having already solved the hard problem.

**Keywords:** meta-problem of consciousness; the hard problem; mind-body problem; phenomenal consciousness

Having already done much to draw attention to the now-famous "hard problem" of consciousness (i.e. the problem of why certain physical events taking place in our brains give rise to phenomenal consciousness), David Chalmers has recently posed yet another problem for philosophers to sink their teeth into: the so-called *meta*-problem of consciousness. This is the problem of explaining why we (or at least many of us) are inclined to think that there *is* a hard problem of consciousness, and to express (or be disposed to express) this intuition through "*problem reports*", in which we say things like "There is a hard problem of consciousness", "It is hard to see how consciousness could be physical", or "Explaining behaviour does not explain consciousness" (Chalmers 2018, 7).

Luckily, this problem is (in Chalmers' view) an *easy* one, in that to solve it we need only provide a physical or functional account of how we come to produce such utterances. Chalmers suggests that we can reasonably expect to be able to formulate such an account, "at

least if we accept that all human behaviour can be explained in physical and functional terms” (Chalmers 2018, 8). While one might wonder whether we *should* accept this assumption, the relevant point for our purposes is that the meta-problem differs from the hard problem in that whereas the latter concerns phenomenal consciousness itself, the former concerns certain judgments and statements that we make *about* phenomenal consciousness. Insofar, then, as judgments and verbal reports in general seem more capable of being explained in physical or functional terms than phenomenal consciousness itself, we have reason to view the meta-problem as more susceptible to physicalist or functionalist solutions than the hard problem.

This feature of the meta-problem is of special interest to Chalmers (2018, 8), as he thinks that “[w]e can reasonably hope that a solution to the meta-problem will shed significant light on the hard problem”, either by “solv[ing] or dissolv[ing]” it completely, or at the very least, by “constrain[ing] the form of a solution” to it. If this is so, then the fact that the meta-problem is easy is doubly welcome, as solving it will help us make progress on a problem that is (or at least appears) significantly harder. It is here that my basic disagreement with Chalmers lies. While I agree that the meta-problem constrains the hard problem, in that any feasible solution to the hard problem ought to be able to provide some account of why we make problem reports, I am sceptical of the idea that these constraints will be such as to “shed significant light on the hard problem.”<sup>1</sup>

My scepticism is owing to the fact that, as Chalmers defines it, the meta-problem seems to admit of a number of different solutions, and when confronted with an individual who emits the kinds of problem reports on which the meta-problem is based, it may be difficult if not impossible to tell which of these solutions gives the *correct* explanation of why *this* individual is disposed to emit such reports (particularly when considering individuals, such as extra-

---

<sup>1</sup> A great deal depends, of course, on what one means by “shed significant light on.” My criticisms apply only to interpretations according to which “shedding significant light on” the hard problem requires ruling out at least one of the general stances on it discussed below. It may be that Chalmers means something much weaker than this, in which case he might accept the bulk of the following discussion.

terrestrials or future AIs, who are physically very different from us). This may apply to our own case as well, as we may have no way (or at least no *easy* way) of determining which among the various possible solutions to the meta-problem correctly explains why we make the problem reports that we do.

This indeterminacy poses a problem for those like Chalmers who expect work on the meta-problem to shed light on the hard problem, because the various solutions to the meta-problem reflect differing stances on the hard problem itself. Thus, if we cannot tell which of these solutions applies to any particular case, then we will be unable to use evidence about problem reports to determine which stance on the hard problem is most credible. The most we can do is rule out those stances on the hard problem that cannot provide *any* viable account of how problem reports are produced. This constraint, however, does little to winnow down the field of competitors, for most people, I take it, are interested in the hard problem primarily because they are interested in knowing which of the following three general solutions<sup>2</sup> to the problem is correct:

- *Dualism/Non-reductive realism* – Consciousness exists and cannot be identified with the physical events that give rise to it or any function that those events perform. Such events give rise to conscious experience by virtue of certain brute correlations between conscious experiences and physical events or functional roles that cannot be explained in purely physical or functional terms.

---

<sup>2</sup> Some might prefer to describe these as “reactions” to the hard problem, reserving the term “solution” for accounts that explain why certain physical events give rise to the particular phenomenal states that they do *without* appealing to brute, inexplicable psychophysical correlations. This would in effect be to say that any solution to the hard problem must be reductive realist in form; on this view, dualism/non-reductive realism and eliminativism are merely ways of claiming that the hard problem is unsolvable or illusory. Those who prefer this way of putting things should substitute “reaction” for “solution” throughout. Put in these terms, my basic contention is that while Chalmers’ meta-problem might be useful in ruling out certain kinds of solutions to the hard problem (when the term “solution” is restricted to reductive realist accounts), this kind of test will not enable us to rule out any of the three main reactions to the hard problem entirely, and (importantly) it is the choice between these three *reactions* that is of primary interest to us in addressing the hard problem. This latter point is the reason why I think “shedding significant light on” the hard problem requires ruling at least one of these reactions out.

- *Reductive realism* – Consciousness *just is* certain physical events taking place in our brains, or a complex function that is realized by those events. There is a physical or functional explanation for why the physical events that realize or are identical to a given experience have the features of that experience as opposed to some other experience or none at all.<sup>3</sup>
- *Eliminativism* – There is no consciousness for the physical events taking place in our brains to give rise to, for consciousness doesn't exist.

Each of these stances on the hard problem seems capable of providing a viable account of how problem reports are produced. Consequently, if we are unable to tell which of those accounts applies in any given case (including our own), then the meta-problem won't help us to decide between these three stances on the hard problem. With respect to the hard problem, we'll thus be left in more or less the same position as we were before.

Chalmers divides potential solutions to the meta-problem into two classes: *realist* and *illusionist*. *Realist* solutions assume that phenomenal consciousness exists and has many of the distinctive features that we are inclined to ascribe to it (e.g. that “consciousness is primitive and non-physical” and “cannot be physically explained”) (Chalmers 2018, 43). *Illusionist* solutions, in contrast, hold either that consciousness does *not* exist (a position Chalmers calls *strong* illusionism), or else that it exists but lacks many of the features that it seems to us to have (a position Chalmers calls *weak* illusionism). These three types of solutions to the meta-problem correspond to the three general stances on the hard problem distinguished above. Those who adopt a dualist/non-reductive realist stance on the hard problem will thus seek to provide a realist response to the meta-problem, eliminativists will stand in need of a strong

---

<sup>3</sup> Reductive realists might hold differing views on whether this explanation is accessible to us. More on this below.

illusionist solution<sup>4</sup>, and reductive realists will look to solve the meta-problem in weak illusionist terms.

Chalmers seems to grant that while all three of these positions – realism, weak illusionism, and strong illusionism – have their difficulties, each can nevertheless in principle provide an explanation of how our problem reports are produced. In each case he envisions the relevant account as involving the specification of a “meta-problem process”, wherein certain higher-order introspective models “attribute special mental states...to ourselves when our brains are in certain lower-order cognitive states (such as perception, attention, or access-consciousness...)” (Chalmers 2018, 40-1). Realists have the option of holding either that conscious states correlate with the states that realize these meta-problem processes, or that conscious states themselves realize certain of the functional roles that these processes involve.<sup>5</sup> Weak illusionists can either identify consciousness with the lower-order cognitive states that are targeted by meta-problem processes, or with the higher-order introspective models that attribute special mental states to ourselves when we are in such lower-order states. Strong illusionists will of course refrain from identifying consciousness with either the higher- or lower-order states involved in the meta-problem processes they postulate, since they deny that consciousness exists. They may nevertheless seek to explain why consciousness *seems* to exist by identifying the things we erroneously take to be conscious experiences “with the special primitive properties that are (or seem to be) attributed [to us] by our [higher-order] introspective models”, even though “[n]o such special primitive properties are instantiated in our brains” (Chalmers 2018, 43).

---

<sup>4</sup> Frankish (2016, 11-2, 21-2) is reluctant to associate the strong illusionist view he defends with eliminativism, as the latter term carries certain connotations that he sees as inessential to strong illusionism. Here I am interested in eliminativism solely as the view that phenomenal consciousness does not exist. I associate the two positions (eliminativism and strong illusionism) solely on the basis of their shared commitment to this thesis.

<sup>5</sup> Chalmers (2018, 41) notes that realists also have the option of holding that the meta-problem simply has no solution. I set this option aside, as I’m interested in the *solutions* that realism, weak illusionism, and strong illusionism can provide to the meta-problem.

Regardless of the general stance one takes on the hard problem, one thus has access to a potential solution to the meta-problem. Eliminativists can appeal to the strong illusionist solution, reductive realists can choose between either of the weak illusionist options, and dualists/non-reductive realists can avail themselves of either of the realist solutions. If we can't find a way of deciding between these solutions, then we seem likely to end up in a three-way stalemate that will prevent us from ruling out certain stances on the hard problem on the grounds that they are unable to provide an effective solution to the meta-problem. In order to leverage the meta-problem into a means of making the hard problem easier, we will thus need to show that some of the aforementioned solutions are not worthy of serious consideration.

Of the realist, weak illusionist, and strong illusionist solutions outlined above, Chalmers finds strong illusionism and forms of realism that ascribe consciousness a functional role in the production of problem reports to be the most promising. This is enticing, for if we have sufficient grounds for leaving weak illusionism by the wayside, we might rule out the associated reductive realist stance on the hard problem as well. However, Chalmers' objections to weak illusionism strike me as themselves a bit weak. His basic criticism seems to be that weak illusionism doesn't hold any promise of resolving the hard problem, for the hard problem doesn't, in his view, depend on consciousness actually possessing any of the features that weak illusionists refuse to grant it. As Chalmers (2018, 49) puts it:

[T]he hard problem does not turn on the claim that consciousness is intrinsic, or non-physical, or non-representational, or primitive, and so on. For example, we can be agnostic about whether consciousness is intrinsic, or hold that it is extrinsic, and the hard problem arises as strongly as ever... The same goes for non-physicality, non-representationality, primitiveness, ineffability, and so on.

Having granted that consciousness exists, the weak illusionist is thus unable, in Chalmers' view, to provide any solution to the hard problem, for the problem persists even if consciousness lacks the features that weak illusionists claim it lacks.

I find this objection puzzling for two reasons. First, Chalmers' criticism of weak illusionism is premised on the idea that we should prefer those solutions to the meta-problem that have the potential to give us a solution to the hard problem as well. This seems sensible enough, but it assumes something that Chalmers has yet to demonstrate: viz. that solving the meta-problem *will* in fact shed significant light on the hard problem. Couldn't it be that weak illusionism gives the correct account of how problem reports are produced (at least in some cases) even if this gives no answer to the question of why certain of the physical or functional states involved in this account are conscious experiences? We can, of course, *guarantee* that solving the meta-problem will render the hard problem more tractable by refusing to countenance any solutions to the former that fail to shed light on the latter. But this just assumes from the outset that the correct account of how our problem reports are produced *will* in fact explain why certain physical events in our brains give rise to (or *seem* to give rise to) the conscious experiences that they do.

Against this assumption, weak illusionists might reasonably hold that in the course of explaining why we produce problem reports and make various false statements and judgments about the nature of our phenomenal states, an adequate theory of phenomenal consciousness will also explain why there is (or at least *appears* to be) an unbridgeable explanatory gap between such states and the physical or functional states with which they are identical. In support of this, weak illusionists might use the phenomenal concept strategy to argue that this gap is due merely to certain differences between the phenomenal and physical/functional concepts that we apply to our brain states, and is thus purely epistemic in nature (Papineau

2002; Hill 1997; Loar 1990).<sup>6</sup> Different stances might then be taken on the question of whether this epistemic gap can be bridged *by us*, with some weak illusionists (e.g. Carruthers 2000) holding that the explanation for why a physical event taking place in a person's brain is a conscious experience (or is identical to one kind of conscious experience as opposed to another) is at least in principle accessible to us, while others (e.g. McGinn 1989) hold that while some such explanation exists, it is nevertheless "cognitively closed" to us, due to the fact that the differences between our phenomenal and physical/functional concepts that generate the explanatory gap are an ineluctable part of our cognitive makeup.<sup>7</sup> Still others (e.g. Papineau 2002) might hold that while no explanation can be given for why a physical event taking place in a person's brain is the conscious experience that it is, this is merely because identities cannot be explained; the assumption that such an explanation is called for derives, on this view, from the mistaken intuition that phenomenal and physical states are distinct, which is again attributable to certain differences between our phenomenal and physical/functional concepts.

Weak illusionists who adopt either of the latter two views might hold that even if we can't solve the hard problem by explaining why certain physical events are the conscious experiences that they are, there are nevertheless compelling reasons for embracing reductive realism (e.g. the difficulties that dualists/non-reductive realists face in accounting for the causal efficacy of phenomenal states, or that eliminativists have in explaining away introspective

---

<sup>6</sup> Chalmers (2018, 21-2; 2007) and Frankish (2016, 25-6) both criticize the phenomenal concept strategy on the grounds that any account of phenomenal concepts will make such concepts either (a) too thin to explain our problem intuitions and qualify as genuinely phenomenal, or (b) too robust to be explainable in a way that is compatible with reductive realism. Balog (2012), Papineau (2007, 136-43), and Carruthers and Veillet (2007) provide responses to Chalmers that seem to apply to Frankish's objections as well.

<sup>7</sup> Carruthers (2000, 62-4) explicitly rejects this second option. Papineau (2002, 178, 197) seems to as well. I'm unsure whether McGinn would associate his view with the phenomenal concept strategy in the manner suggested above. My intention, however, is merely to note one potential position that weak illusionists might adopt, which agrees with McGinn in viewing the explanatory gap as both purely epistemic in nature and as something that *we* may be unable to bridge, due to certain constraints on how we think. (The central idea is expressed in McGinn's (1989, 361-2) claim that "the nature of the psychophysical connection has a full and non-mysterious explanation in a certain science, but... this science is inaccessible to us as a matter of principle.") Another advocate of this approach might be found in Levine (2001), whose combined endorsement of materialism and "modest qualophilia" seems to commit him to something along these lines.



evidence for the existence of conscious experiences).<sup>8</sup> Insofar as this kind of position is available to weak illusionists, they might willingly concede that their solution to the meta-problem doesn't explain why a physical event taking place in a person's brain is a conscious experience (or is identical to one kind of conscious experience as opposed to another), so long as a plausible explanation can be given for why no such explanation is accessible to us (Balog 2012, 16-21; Papineau 2002, chaps.5-7).

Second, as Chalmers points out, weak illusionists are apt to deny that conscious experiences are non-physical, even if they *appear* to be so. He claims, however, that even if weak illusionists are right about this, the hard problem will remain unresolved. As he puts it: “[I]f the appearance that consciousness is non-physical is an illusion, then consciousness is physical, and the letter of materialism is saved. But this does little to address the hard problem: we still have no explanation of why there is something it is like to be us” (Chalmers 2018, 49). It seems to me, however, that barring things like identity statements and fundamental physical facts that cannot be explained in terms of any more basic facts, for something to *be* physical is (at least in part) for it to be in principle explainable in purely physical terms. Thus, if weak illusionists are right to claim that consciousness is physical, then while there may be no explanation for why a given conscious experience C is the physical event P that it is (since identities cannot be explained), there must at least be a physical explanation for why P has the various features possessed by C whereas other physical events do not, even if that explanation is for some reason inaccessible to us.

Questions may of course be raised as to whether weak illusionists are *justified* in claiming that consciousness is physical and hence physically explainable in this way, especially if the relevant explanation is said to be cognitively closed to us. Weak illusionists who endorse

---

<sup>8</sup> Levine (2001, chaps.1, 5) argues against dualism/non-reductive realism and eliminativism on these grounds. Frankish (2016, 25) also draws attention to the dualist/non-reductive (or, in his terms, “radical”) realist’s problems in avoiding the “threat of epiphenomenalism.” Papineau (2002, chap.1) and Carruthers (2000, 2-3) reject dualism/non-reductive realism for similar reasons.

the latter claim may have no way of answering such questions except by appealing to the kinds of negative arguments mentioned above, which rest the case for reductive realism on the problems facing its dualist/non-reductive and eliminativist rivals. Whether or not such arguments are enough to justify belief in the existence of a reductive realist solution to the hard problem that we have no epistemic access to, the point remains that if conscious experiences *are* identical to certain physical states involved in producing our problem reports as weak illusionists claim, then there must be a physical explanation for why those physical states have the features possessed by the conscious experiences that they are identical to (albeit, again, perhaps not one that we can comprehend). Weak illusionism thus can't be dismissed on the grounds that it offers no solution to the hard problem, for if weak illusionists are right to deny that consciousness is non-physical, then the hard problem *must* be solvable in purely physical terms (even if that solution lies beyond our epistemic reach). While Chalmers (2018, 49, emphasis added) is therefore right that on versions of weak illusionism that treat the solution to the hard problem as cognitively closed to us “*we* still have no explanation of why there is something it is like to be us,” so long as weak illusionists can provide reasons for thinking that some reductive realist explanation of this fact exists while also explaining why we lack access to this explanation (if indeed we do), it seems unfair to treat the view as failing to address the hard problem.

Chalmers isn't alone in viewing weak illusionism as the weakest of the three available solutions to the meta-problem. Frankish (2016; 2012), who unlike Chalmers advocates strong illusionism, seems to share this view as well.<sup>9</sup> Frankish's basic objection to weak illusionism is that it is unable to avoid collapsing into either strong illusionism or realism. For in order to distinguish their view from realism and render phenomenal consciousness susceptible to

---

<sup>9</sup> Frankish's criticisms are actually focused on what I'm calling the reductive realist solution to the hard problem, but I expect he would see these as carrying over to the associated weak illusionist solution to the meta-problem as well.

physical or functional explanation, weak illusionists must deny that phenomenal states have certain features (e.g. intrinsicity, ineffability, and subjectivity) that realists typically ascribe to them. But, Frankish argues, once we strip phenomenal states of these features, there is no longer any reason to describe such states as genuinely phenomenal, and thus nothing to distinguish weak illusionism from strong illusionism. “How could a *phenomenal residue* remain,” Frankish (2012, 669) asks, “when intrinsicity, ineffability, and subjectivity have been stripped away?” Frankish (2016, 16) thus calls on weak illusionists to “face up to the challenge of articulating a concept of the phenomenal that is both stronger than that [provided by the strong illusionist] and weak enough to yield to [physical or functional] treatment.” Regarding the prospects for meeting this challenge, Frankish is quite frank: “I doubt this is possible.”

I’m not so certain, however, that this challenge is as daunting as Frankish suggests. It seems to me that weak illusionists might pick out phenomenal states demonstratively (as *that* sensation, or *that* kind of smell), and conceptualize them as those felt qualities that can be referred to through such acts of introspective demonstration, thereby leaving it open whether the phenomenal states thus referred to also have the additional features (e.g. intrinsicity, ineffability, and subjectivity) that realists ascribe to them.<sup>10</sup> Frankish (2012, 670-2) considers and rejects this option on the grounds that introspective demonstration may fail to secure a common referent for phenomenal concepts across individuals who hold different theories of phenomenal consciousness, thereby leaving us without any clear sense of what the nature of the phenomenal states that we’re supposedly referring to is, or indeed whether they qualify as genuinely phenomenal (as the weak illusionist claims). He notes, e.g., that whereas “anti-transparentists” (e.g. Block 1990 and Peacocke 1983) will take such introspective

---

<sup>10</sup> Such an account could fit well with the constitutional accounts of phenomenal concepts endorsed by Balog (2012) and Papineau (2007; 2002), which treat such concepts as involving an instance of the phenomenal states they refer to.

demonstrations to pick out intrinsic properties of experiences, “transparentists” (e.g. Dretske 1995, Tye 1995, and Harman 1990) will treat them as referring instead to properties of external objects that our experiences represent.

The obvious response to this is that individuals who hold such differing views on the nature of the states they are introspectively demonstrating may nevertheless still be picking out the same properties despite their disagreement about what those properties are. Against this, however, Frankish notes that if transparentists and anti-transparentists are in fact “attending to” the same phenomenal properties, then at least one party must be “radically mistaken” about the egocentric location of those properties. He then infers from this that in attempting to form demonstrative thoughts about phenomenal states, at least one of party “will not succeed in thinking a demonstrative thought at all,” since (following Evans 1982, chap.6), he holds that “to identify a spatio-temporal particular demonstratively one must be able to locate it and track it in egocentric space.”

There are, I think, a few different ways that weak illusionists might respond to this worry. First, in order to derive the conclusion that either transparentists or anti-transparentists fail to accurately locate phenomenal properties in egocentric space, Frankish (2012, 671, my emphasis) must grant that they are indeed “*attending to* properties of the same general type” but ascribing different locations to them. Why then can’t the fact that they are attending to the same kinds of properties (while disagreeing about their location) suffice to ensure that the demonstratives they direct towards the properties they are *both* attending to refer to the same thing (viz. certain phenomenal states)? Second, even if the introspective demonstratives used by certain parties within the weak illusionist camp don’t refer to any genuinely phenomenal properties (or indeed anything at all), I don’t see why this poses a problem for weak illusionism in general. So long as there is *some* concept of the phenomenal available that enables us to pick out phenomenal states without committing ourselves on the question of whether they have the

various features that realists ascribe to them, why should it matter if *some* weak illusionists incorrectly question the coherence of that concept and instead advocate some alternative concept of the phenomenal that fails to refer to anything genuinely phenomenal (or anything at all)?

Lastly, if confronted with doubts as to whether they're picking out any genuinely phenomenal properties, it seems to me that weak illusionists of any stripe (be they transparentists or anti-transparentists) can simply piggy-back the reference of their proposed concept of the phenomenal on the introspective demonstratives of realists by saying something like: "Whatever properties you realists are referring to when you demonstratively identify the intrinsic, ineffable, subjective qualities that you speak of, we're referring to the same thing when we speak of phenomenal properties; we're just denying that those properties are indeed intrinsic, ineffable, and subjective in the way that you claim." If weak illusionists can anchor the reference of their concept of the phenomenal to that of the introspective demonstratives employed by realists in this way without taking on board the realists' commitment to viewing the properties thus picked out as intrinsic, ineffable, and subjective, then they seem to have a way of talking about genuinely phenomenal states that leaves room for debate over whether such states really have the various features that realists ascribe to them. For these reasons, I don't think that weak illusionism can be dismissed for failing to articulate a concept of the phenomenal that is genuinely phenomenal while also leaving it open whether phenomenal states are actually intrinsic, ineffable, and subjective. Despite Frankish's scepticism, it seems to me that such a concept can be derived from the application of introspective demonstratives to our own phenomenal states.

With weak illusionism back on the table, we seem to be back where we started with respect to the hard problem, facing a choice between dualism/non-reductive realism, reductive realism, and eliminativism. Perhaps progress can be made by instead questioning the realist

and strong illusionist solutions to the meta-problem associated, respectively, with dualism/non-reductive realism and eliminativism. While Chalmers finds these to be the most promising solutions available, he notes that each still faces certain difficulties. If these difficulties turn out to be serious enough to cast doubt on the viability of realism or strong illusionism, we might use this as grounds for rejecting the associated dualist/non-reductive realist or eliminativist stances on the hard problem, thereby vindicating Chalmers' suggestion that focusing on the meta-problem will make the hard problem more tractable. However, I don't think the difficulties that Chalmers raises are serious enough to treat either realism or strong illusionism as less worthy of consideration than its rivals.

Chalmers' main objection to strong illusionism is that it is incompatible with the seemingly obvious Moorean fact that people sometimes feel pain. However, just like Moore in his proof of an external world, here too Chalmers may be accused of begging the question against his strong illusionist opponents, who will naturally challenge Chalmers' entitlement to the claim that people do sometimes feel pain. After all, the strong illusionist *denies* this, and moreover has an account of why it nevertheless *seems* obvious to us that people sometimes feel pain. It's therefore a bit unfair to use the claim that people feel pain in an argument *against* strong illusionism. Following Frankish (2016, 27-8), strong illusionists might offer the following argument as grounds for thinking that Moorean intuitions about the existence of phenomenal states like pain are in fact mistaken:

1. When we encounter any seemingly anomalous phenomenon like phenomenal consciousness that "resists explanation in physical terms or is detectable only from a certain perspective,... the simplest explanation is that it is illusory."
2. Consequently, "if there is even a remote possibility that we are mistaken about the existence of phenomenal consciousness, then there is a strong abductive inference to the conclusion that we are in fact mistaken about it."

3. There is, however, a possibility that we are so mistaken, for in order for our phenomenal states to have any impact on our mental lives, we must form introspective mental representations of them, yet “we have no introspective way of checking the accuracy of our introspective representations, and so cannot rule out the possibility that they are non-veridical.”<sup>11</sup>
4. Therefore, there is a strong abductive inference to the conclusion that our Moorean intuitions about the existence of phenomenal states are mistaken.

So long as this argument remains available to strong illusionists, advocates of the view seem justified in treating the Moorean facts that Chalmers appeals to as in truth nothing more than faulty intuitions. Though I sympathize with Chalmers’ stance on this issue, any argument against illusionism of the sort he provides consequently seems destined to end in a stalemate (as his imagined dialogue between the realist and the strong illusionist on pp.54-5 illustrates). I therefore don’t think we can take strong illusionism’s incompatibility with Moorean common sense as grounds for rejecting it.

With respect to realism, I share Chalmers’ opinion that the most plausible realist solutions to the meta-problem will be those that treat consciousness as realizing certain of the states involved in producing our problem reports. Such views have the advantage of assigning consciousness an actual functional role in the production of our reports about it. Chalmers’ worry, however, is that the relation between consciousness and the role that it is assigned on such accounts will turn out to be too contingent, as one might naturally ask what reason we have to assume in any given case that the role is (or must be) performed by a conscious experience, rather than by some non-conscious physical state.

---

<sup>11</sup> Frankish (2016, 28) further notes that “in so far as we can check [the accuracy of our introspective representations of phenomenal states], through external inspection of our brain states, they appear to be non-veridical; the properties represented do not show up from other perspectives.”

There are, however, at least two ways in which realists might address this worry. As Chalmers (2018, 48-9) notes, one option would be to claim (a) that *only* conscious experiences have the causal powers necessary to realize certain of the states involved in producing problem reports, and (b) that the causal powers that enable conscious experiences to play this role are also *essential* to them.<sup>12</sup> This would imply that conscious experiences cannot occur *without* being in a position to produce problem reports, and moreover that such reports cannot be produced *without* conscious experiences. Justification for (b) might be derived from the adoption of a more general dispositional essentialist theory of properties, according to which properties (e.g. the property of having a certain kind of conscious experience) bestow the same causal powers on their bearers in all possible worlds. While it's less clear what reasons realists might give in support of (a), I don't think realists need to endorse this claim anyway. Commitment to (a) would seem necessary if any adequate realist solution to the meta-problem had to rule out all other solutions, by entailing that problem reports couldn't be produced in the absence of consciousness. But this seems far too strict a requirement on what a realist solution to the meta-problem must show.<sup>13</sup> Surely it's enough if the realist can make a plausible case that problem reports are at least *sometimes* produced by processes involving non-physical conscious experiences. If this is so, then realists can allow that while non-physical conscious experiences play a role in the production of some problem reports, such reports could also be (and perhaps on occasion *are*) produced by meta-problem processes consisting entirely of

---

<sup>12</sup> Chalmers cites Mørch (2018) as advocating a position of this sort, although she is concerned not with problem reports but rather with the avoidance and pursuing behavior produced by experiences of pain and pleasure. It's unclear whether Mørch would accept (a). The closest she comes to affirming something like it is on p.309, where she states that "it is not clear whether there are any physical powerful qualities with the same explanatory features as phenomenal powerful qualities." But there she is contrasting the *epistemically* contingent connection between physical qualities and their powers with the kind of "intelligible connection" that she posits between experiences of pain or pleasure and avoidance or pursuit (which she suggests is "not revealed by physics or the physical sciences, but rather only by first-person experience") (p.310). This is compatible with the thesis that the avoidance and pursuing behavior produced by pain and pleasure could be produced by other, non-conscious means as well.

<sup>13</sup> The same cannot be said of strong illusionist solutions, which must insist that problem reports are never produced in the way that realists or weak illusionists suggest in order to maintain consistency with their associated eliminativist stance on the hard problem. Weak illusionists must likewise deny that problem reports are ever produced in the way that realists suggest in order to maintain consistency with reductive realism.



physical states. Some realists might, e.g., wish to take this kind of position with regard to any problem reports produced by future AIs or extra-terrestrials with bodies very different from our own. One might reasonably hold that we simply cannot know whether non-physical conscious experiences play a role in producing the problem reports of such beings or not.

Chalmers (2018, 44-9) worries that this leaves the realist vulnerable to “debunking” arguments for illusionism, which ask what reason we have to think that consciousness exists and is non-physical if our problem reports could just as easily have been produced by processes consisting entirely of physical states. This might pose a difficulty if problem reports were the *only* indication we had that consciousness exists and is non-physical. Yet surely there are other reasons for thinking that consciousness exists and is non-physical besides the existence of problem reports! Introspective reflection on conscious experiences and the conspicuous absence of any widely agreed-upon physical or functional account of phenomenal consciousness come to mind as other sources of support for this view. Realists can thus allow that problem reports might sometimes be produced by processes in which consciousness plays no part without thereby relinquishing the only grounds for thinking that consciousness exists and is non-physical.

Moreover, even if problem reports *were* the only evidence for the view that consciousness exists and is non-physical, I’m still not sure that realists must deny that such reports could be produced by purely physical processes, for I don’t think the debunking arguments that Chalmers claims realists would otherwise expose themselves to are as threatening as Chalmers makes them out to be. The fact that the evidence for a given theory can also be explained by another rival theory is in itself no reason to think that the former theory and the explanation it gives for that evidence are false. It just means that the choice between those two theories is underdetermined by that particular body of evidence. This, I suggest, is the situation with respect to problem reports and the rival explanations that realist

and illusionist solutions to the meta-problem provide for them. Even if such reports were the primary or *only* source of evidence for the view that consciousness exists and is non-physical, the fact that they *could* be produced by purely physical processes does not by itself imply that they are never produced by processes in which non-physical conscious experiences play a functional role. The worries that Chalmers raises about the viability of realist solutions to the meta-problem thus strike me as unfounded.

Based on the foregoing discussion, it appears that none of the three main types of solution to the meta-problem is seriously threatened by the difficulties that Chalmers raises for it. We consequently seem left with a three-way stalemate between realism, weak illusionism, and strong illusionism. Our hopes of using the meta-problem to make progress on the hard problem by ruling out those stances on the hard problem that cannot provide a viable account of how problem reports are produced have thus come to no avail. For all three main stances on the hard problem seem capable of fielding viable candidate solutions to the meta-problem, and we have no obvious means of deciding which of these solutions gives the *correct* account of how any given problem report is produced without presupposing one of the three stances on the hard problem. While Chalmers has performed a welcome service to philosophy by drawing attention to another important and interesting problem of consciousness, I'm therefore sceptical of the idea that work on the meta-problem will make the hard problem any easier.

## Works Cited

- Balog, K. 2012. "In Defense of the Phenomenal Concept Strategy." *Philosophy and Phenomenological Research* 84: 1–23.
- Block, N. 1990. "Inverted Earth." *Philosophical Perspectives* 4: 53–79.
- Carruthers, P. 2000. *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge University Press.
- Carruthers, P. and Veillet, B. 2007. "The Phenomenal Concept Strategy." *Journal of Consciousness Studies* 14: 212–36.
- Chalmers, D. 2007. "Phenomenal Concepts and the Explanatory Gap." In *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, eds. Alter and Walter. Oxford University Press.
- . 2018. "The Meta-Problem of Consciousness." *Journal of Consciousness Studies* 25: 6–61.
- Dretske, F. 1995. *Naturalizing the Mind*. MIT Press.
- Evans, G. 1982. *The Varieties of Reference*. Oxford University Press.
- Frankish, K. 2012. "Quining Diet Qualia." *Consciousness and Cognition* 21: 667–76.
- . 2016. "Illusionism as a Theory of Consciousness." *Journal of Consciousness Studies* 23: 11–39.
- Harman, G. 1990. "The Intrinsic Quality of Experience." *Philosophical Perspectives* 4: 31–52.
- Hill, C. 1997. "Imaginability, Conceivability, Possibility and the Mind-Body Problem." *Philosophical Studies* 87: 61–85.
- Levine, J. 2001. *Purple Haze: The Puzzle of Consciousness*. Oxford University Press.
- Loar, B. 1990. "Phenomenal States." *Philosophical Perspectives* 4: 81–108.
- McGinn, C. 1989. "Can We Solve the Mind-Body Problem?" *Mind* 98: 349–66.
- Mørch, H. 2018. "The Evolutionary Argument for Phenomenal Powers." *Philosophical Perspectives* 32: 293–316.
- Papineau, D. 2002. *Thinking About Consciousness*. Oxford University Press.
- . 2007. "Phenomenal and Perceptual Concepts." In *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, eds. Alter and Walter. Oxford University Press.
- Peacocke, C. 1983. *Sense and Content*. Oxford University Press.
- Tye, M. 1995. *Ten Problems of Consciousness*. MIT Press.