

Graham Peebles¹

Two Problems for Non-Inferentialist Views of the Meta-Problem

Abstract: *The meta-problem of consciousness is to explain why we think that there is a hard problem of consciousness. On Chalmers' view of the meta-problem, our judgments about the hard problem of consciousness arise non-inferentially as a result of introspection. I raise two problems for such a non-inferentialist view of the meta-problem. It does not seem to match the psychological facts about how we come to the realization of the hard problem, and it is unclear how the view can bridge the gap between the content of introspection and the content involved in formulations of the hard problem. The inferentialist view of the meta-problem, on which the hard problem results from inference, explains both the psychology and content introduction. We should therefore prefer an inferentialist view of the meta-problem.*

1. The Meta-Problem, Introspection, and Inference

The meta-problem of consciousness is to explain the origin of the hard problem of consciousness, why a physical explanation of what is perhaps best described as the 'feel' of conscious experience seems to be in principle elusive. To make progress on the hard problem we

Correspondence:

Email: peebles@ams.eng.osaka-u.ac.jp

¹ Intelligent Robotics Laboratory, Graduate School of Engineering Science, Osaka University, Japan.

should therefore look at the meta-problem. Chalmers conceives of the meta-problem as follows. The hard problem arises from introspection of conscious experience. What need explaining are our ‘problem intuitions’ which are dispositions to make ‘problem judgments’ about experiences, which we express through ‘problem reports’. Here are some examples of problem reports that Chalmers provides: “‘There is a hard problem of consciousness”, “It is hard to see how consciousness could be physical”, “Explaining behaviour does not explain consciousness”, and so on’ (Chalmers, 2018, p. 7). According to Chalmers, these intuitions reflect our underlying sense that consciousness provides us with a unique explanatory problem. They reflect ‘our sense that there is some sort of special problem involving consciousness, and especially some sort of gap between physical processes and consciousness. For example, “I can’t see how consciousness could be physical” is a problem report, and the disposition to judge and report this is a problem intuition’ (*ibid.*, p. 12). This has the surprising effect of making the meta-problem one of the easy problems, and therefore a problem of explaining the underlying mechanisms responsible for the production of the problem reports: ‘At least if we accept that all human behaviour can be explained in physical and functional terms, then we should accept that problem reports can be explained in physical and functional terms. For example, they might be explained in terms of neural or computational mechanisms that generate the reports’ (*ibid.*, p. 8).

Such a view is subject to various deeper elaborations, from a causal relation between the introspected properties of experience and the problem judgments, to a view on which consciousness is the matter or vehicle from which mental states in the process are constituted. I would like to raise two problems for this general view of the meta-problem, which I will call ‘non-inferentialism’. These problems suggest that the alternative view of the meta-problem, ‘inferentialism’, something that Chalmers forecloses towards the beginning of the paper, is the better one. Here is Chalmers explicitly rejecting the view that I would like to defend.

Problem intuitions can result from inferences. So that judgments that result from philosophical arguments will count as problem intuitions. Still, it is plausible that, in solving the meta-problem, the most important problem intuitions will be non-inferential judgments that arise prior to philosophical argument, and I will focus especially on judgments of this sort. (*ibid.*, pp. 11–12)

We have two competing views of the meta-problem and the way that the hard problem arises: that the problem judgments arise non-inferentially, and that they arise inferentially. If the latter view is correct, then the meta-problem submits to a philosophical explanation: the source of the hard problem is to be found in reasoning. Chalmers favours the alternative view, or at least what is effectively the alternative view, that the most important and theoretically central problem judgments originate non-inferentially as a result of introspection.

The problems that I would like to raise for the non-inferential view concern its psychological plausibility and its prospects of explaining the origins of the content of the problem judgments. I will draw two conclusions. First, that non-inferentialism is not psychologically plausible. It is not psychologically plausible that the problem judgments are arrived at through a non-inferential process. Second, that it is unclear how such a process that non-inferentially prompts judgments strictly from introspection could explain the introduction of the content required to formulate the hard problem.

2. Two Problems for Non-Inferentialism

I ask the reader to consider the circumstances in which the hard problem first struck them. Presumably they pondered the problem for some time, perhaps relatively briefly, perhaps at length, before satisfying themselves that they understood it — and then likely pondered at much greater length over the soundness of the argument. What, I submit, did not happen was that they introspected experience and were then struck immediately — non-inferentially — by the hard problem. This would be an extremely odd experience for one to undergo. This would involve a quite jarring juxtaposition of two very different states with very different contents — a juxtaposition that I, at least, do not experience. When one introspects an experience with content *p*, the act of introspection results in a judgment with the content *there is an experience with content p* (e.g. *I am having an experience of a blue bead*). Were the non-inferential view of the meta-problem correct, then this act of introspection should also non-inferentially prompt a problem judgment. However, this seems false as a description of introspection and the formation of the problem judgments. There is no such jarring juxtaposition of judgments with such distinct contents. To put it succinctly, introspection is not revelatory of the hard problem.

Further, if there is such a mechanism of introspection that is revelatory of the hard problem, then in at least many cases (in fact, I would

suggest that it is the majority) it fails to work. People tend not to non-inferentially and immediately form the hard problem judgments upon introspection. The proponent of the non-inferential view must therefore explain why this procedure is so unreliable. A further problem: which of the problem judgments are non-inferentially prompted by introspection in which cases and why? Recall Chalmers' various formulations of the hard problem quoted above (and see the further distinctions below). These are quite different, involving notions which are not only philosophically subtle and complex, but many of them are also non-equivalent. Which problem judgment is one non-inferentially prompted to form on which occasions and why? Are they always the same? If they differ between people or across different introspections, then why? What principled reason could one give as to why one, some, or even all of the problem judgments are prompted in some cases but not in others? Different episodes of introspection do not seem to differ enough in their content to prompt such a diverse range of problem judgments. This is so even if one contrasts, for example, introspecting pain and visual experience. What could it be about these conscious states that gives rise to one problem judgment as opposed to another? The non-inferentialist must explain why, for example, on one occasion one might be prompted by introspection of experience of a blue bead to judge that consciousness cannot be physical, and on another occasion introspection of the experience of a middle C note might prompt one to judge that the mental cannot be explained in behavioural terms. These two formulations are not equivalent. Not all physical explanations are behavioural. What could this explanation be? Why would, for example, introspection of visual experience prompt a judgment about the physical, and introspection of auditory experience prompt a judgment about behaviour, or vice versa, and so on? The most likely response will be that in different acts of introspection we attend to different features of consciousness. Granted, but we are still owed an explanation of the link between the objects of attention and the problem judgments non-inferentially formed on the basis of attending to these objects. Here we find another problem.

The content mismatch between introspective judgments and the problem judgments is significant. Consider the quoted examples of problem judgments: 'There is a hard problem of consciousness', 'It is hard to see how consciousness could be physical', 'Explaining behaviour does not explain consciousness', 'I can't see how consciousness could be physical'. There is a significant content mismatch between the data of introspection and such statements of the hard

problem. To further demonstrate the complexity of the problem judgments and the conceptual chasm between their contents and plausible contents of introspective judgments, consider Chalmers' division of the problem intuitions into the following categories: 'explanatory intuitions' that consciousness is hard to explain, 'metaphysical intuitions' that consciousness is non-physical or fundamental, 'knowledge intuitions' that consciousness provides special subjective knowledge, and 'modal intuitions' that, for example, zombies are possible (Chalmers, 2018, p. 12). Consider, in comparison, what introspection does or could tell us about experience. At one end of the spectrum, proponents of transparency (e.g. Harman, 1990) hold that introspection only tells us that experience is a relation, of some sort yet to be determined, to (putative) external objects. If we move to the other end of the spectrum and hold that we can introspect qualia in their classic form (as described in Dennett, 1993), this relatively enriched content of introspection still falls short of the content of the problem judgments. The non-inferentialist must therefore describe and defend a non-inferential process that could introduce the content required to bridge the gap between the contents of introspection and the contents of the problem judgments. It seems as though there must be a link, otherwise the procedure would appear to be random: if the two contents are not linked, how do we explain the resulting judgments to ourselves?

Introspection does not seem to be revelatory of the hard problem, as one would expect if introspection prompted, through a non-inferential mechanism, the problem judgments. Furthermore, if there is such a mechanism, we are owed explanations of why it mostly fails and how and why it produces a wide range of output contents. This latter problem serves to highlight the problem that the non-inferentialist has in explaining how the process bridges the conceptual gap between the contents of introspection and the problem judgments.²

² It should be noted that the question concerns the procedure which generates the problem judgments, not with the correctness or reliability of this procedure. It would not matter for present purposes, therefore, if a view like Pereboom's (2011) were correct. Pereboom argues that introspection misrepresents properties of experience which are conceptually irreducible to physical properties. On his view, this second-order misrepresentation is compatible with there being no such properties actually instantiated in first-order experience. Likewise, some proponents of 'illusionism' (Frankish, 2016; Kammerer, 2016) argue that we suffer some sort of 'introspective illusion', what we may call an 'internal illusion' as opposed to the usual 'external illusions' such as happens when, for example, I mistake a coat on the hook for a person at the door. Much

3. Responses

The first response to consider is that of simply holding the line with respect to the psychology of realizing the hard problem. Could one just insist that introspection is revelatory of the hard problem? I think that this is not plausible. Most readers will be quite familiar with, for example, students who do not accept that there is a hard problem — even after extensive discussion and encouragement to introspect. Are these cases of the mechanism malfunctioning? But, given that this is not at all uncommon, why is the malfunction so pervasive? Perhaps, however, the characterization of non-inferentialism as the view on which the problem judgments are immediate is something of a straw man. Perhaps, that is, the process that generates the problem judgments is non-inferential but it is not immediate. The contrast between inferential and non-inferential is not an entirely clear one. At one end, we can consider a causal model of non-inferentialism on which introspective judgments cause the problem judgments. Old-fashioned sense-data views of perception serve as a good model here. On these views, sense-data cause perceptual beliefs and we can think of introspective judgments and the problem judgments behaving in the same way. At the other end, we have full-blown deliberative reasoning. An important question is: at which point would a view of the meta-problem cease to be non-inferential? Space precludes a detailed, extensive discussion of this question, but I can raise some significant problems for the non-inferentialist.

What the non-inferentialist must seek to avoid is the conclusion that the problem judgments are a result of inference, especially extended deliberative inference. What room is there for a view which rejects immediacy in order to accommodate the psychological lack of immediacy but remains non-inferentialist? Consider the discussion about whether and how one's ordinary perceptual beliefs are justified

requires to be said about this view — if the appearance/reality distinction can be preserved internally, how to understand the notion of an 'internal illusion', why we should suffer such a pervasive illusion, what could the physical explanation of this be, and so on — but these questions, while closely related, would, with the exception of one view, not settle the current problem. The exception would be that there is an introspective mechanism which systematically attributes properties to experience whose nature is transparently and obviously (apparently) non-physical. This view, however, is of a piece with its sibling view on which introspection attributes the properties correctly to experience.

by experience.³ Many philosophers (e.g. Chudnoff, 2013; Pryor, 2000; Huemer, 2001) hold variations of the general view that there is something about the phenomenology of experiences that leads us to ineluctably accept their corresponding perceptual beliefs. However, an analogous view for the meta-problem should be rejected. There is no sense in which, for example, there is a state of seeming in which the hard problem seems to be true which stands between introspection and the problem judgments. The idea would be that a further state accompanies, or is somehow part of, introspection and that this further state allows us to immediately move from introspective judgment to the hard problem. But I, at any rate, fail to find such a state when I introspect.

Another possible response might be that the problem judgments are inferred only from the contents of introspection. While there would be inference on this view, no independent premises would be involved in inferring the problem judgments. This would preserve the core idea of non-inferentialism, deal with the psychological problem, and also explain the introduction of the content required to state the hard problem. The idea would be that, upon making the introspective judgment that p , one does not immediately form the problem judgment that q . The content q is not part of p and the judgment that q is not caused by the judgment that p . Rather, it takes only some brief reflection on p to come to realize that q . This type of view is quite common, and most proponents of phenomenal concepts would agree to something along these lines (although they would disagree on the level of inference required). Chalmers' view of introspection is as follows.

We have introspective models deploying introspective concepts of our internal states that are largely independent of our physical concepts. These concepts are introspectively opaque, not revealing any of the underlying physical or computational mechanisms. Our perceptual models perceptually attribute primitive perceptual qualities to the world, and our introspective models attribute primitive mental relations to those qualities. We seem to have immediate knowledge that we stand in these primitive mental relations to primitive qualities, and we have the sense of being acquainted with them. (Chalmers, 2018, p. 34)

³ Thankfully, we can set aside the vexed question of whether, on any view, one is justified in believing that there is a hard problem. Theories of justification cross-cut the question of the meta-problem, and one could hold any combination of views without contradiction (even if some of the combinations are strange).

The question is: can we get from something like this to the formulation of the hard problem without the aid of independent premises? This will be difficult because of the conceptual complexity of the notions involved in the various formulations of the hard problem. The formulation of the hard problem involves such notions as behavioural, dispositional, physical, functional, representational, intentional, as well as notions of modality, reduction, explanation, fundamentality, and so on. It seems unlikely to me that such notions can be inferred strictly from the contents of introspection. Such notions are part of a wider, inter-defined and interrelated nexus of notions which are introduced in various different ways. We do not acquire, for example, any modal notions or notions of fundamentality or reduction from introspection. The move from introspective judgments to the problem judgments therefore requires the introduction of independent premises, sourced from outwith introspection. The same goes, I believe, for phenomenal concept views on which there is a logical feature of the concepts which articulates introspection which renders them logically distinct from physical concepts. While a view like this may get us some distance along the road, it cannot get us all the way without the aid of independent premises.

Another alternative would be to appeal to something like unconscious inference. The idea would be that there is a subpersonal system, modelled on System 1 or heuristic or rule-of-thumb reasoning, which produces the problem judgments as a result of introspection. There are cases in which we move from one conscious state to another in a non-inferential way, and perhaps the move from introspective judgment to the problem judgments is like these. Consider coming to recognize someone. When we come to see someone as a particular person, we move from seeing them as having certain features to seeing them as that person. We do not consciously infer that they are that person. There are many cases where we move quickly from one conscious state to another, without the reasoning being either inferential in the classic sense, conscious, or both. For example, consider someone who is superstitious and whenever they see a black cat they believe that they will die soon. Here we have a judgment which is both non-inferentially prompted by a perceptual judgment and which features a significant content mismatch. However, I do not think that these examples serve well as analogies. We can split them into two groups. The first group contains those cases in which the relation between the input states and output states is scrutable. In these cases, we can reconstruct the transition in terms of the relation between the input

and output contents. Recognizing someone is a good example of such a case. The way that a particular person looks, their hair, height, gait, etc. leads us to recognize them as that person. The problem judgments do not fall into this group for the reasons outlined above. To move from the content of introspection to the content of the problem judgments requires the introduction of independent premises. The other group contains cases where the reasoning is not scrutable, where we could not reconstruct it. The majority of unconscious processes are like this. But to put the formulation of the hard problem into this group would be to deny that there is a scrutable connection between the contents of introspection and the problem judgments. This does not seem to be correct. It is because of what introspection tells us about conscious experience that we form the problem judgments. We are not struck mysteriously by the hard problem as we would be on this view.

4. Conclusion

The non-inferential view of the meta-problem is not psychologically plausible and will struggle to explain the relation between the content of introspection and the content of the problem judgments. Taking these problems into account, we should instead conceive of the meta-problem inferentially. That is, we do not merely introspect, and perhaps reason strictly from, experience in order to formulate the hard problem. Rather, we reason *about* experience, bringing in independent premises and concepts sourced from outwith introspection. This picture fits much better our experience of coming to the hard problem. The hard problem is not immediate, because it is a matter of sophisticated and extended inference. This sophisticated and extended inference also explains the introduction of the content required to formulate the hard problem. This shifts the question about the hard problem significantly from Chalmers' non-inferentialist view of the meta-problem. Instead of looking for mechanisms which generate the problem judgments and investigating consciousness's role in these mechanisms, we ought to ask if the reasoning that leads to the hard problem is sound. In my opinion it is not, but that is a question for another day.

Acknowledgments

I would like to thank audiences at the University of Tokyo, University of Osaka, and the Consciousness Research Network 2019 Meeting in

Nagoya, especially Takeshi Akiba, Lok Chi Chan, Aviv Hoffman, John O’Dea, Lu Teng, and Theodore Paradise, for their comments and suggestions, as well as an anonymous referee.

References

- Chalmers, D.J. (2018) The meta-problem of consciousness, *Journal of Consciousness Studies*, **25** (9–10), pp. 6–61.
- Chudnoff, E. (2013) *Intuition*, New York: Oxford University Press.
- Dennett, D. (1993) Quining qualia, in Goldman, A. (ed.) *Readings in Philosophy and Cognitive Science*, pp. 381–414, Cambridge, MA: MIT Press.
- Frankish, K. (2016) Illusionism as a theory of consciousness, *Journal of Consciousness Studies*, **23** (11–12), pp. 11–39. Reprinted in Frankish, K. (ed.) (2017) *Illusionism as a Theory of Consciousness*, Exeter: Imprint Academic.
- Harman, G. (1990) The intrinsic quality of experience, *Philosophical Perspectives*, **4**, pp. 31–52.
- Huemer, M. (2001) *Skepticism and the Veil of Perception*, New York: Rowman and Littlefield.
- Kammerer, F. (2016) The hardest aspect of the illusion problem — and how to solve it, *Journal of Consciousness Studies*, **23** (11–12), pp. 124–139. Reprinted in Frankish, K. (ed.) (2017) *Illusionism as a Theory of Consciousness*, Exeter: Imprint Academic.
- Pereboom, D. (2011) *Consciousness and the Prospects of Physicalism*, Oxford: Oxford University Press.
- Pryor, J. (2000) The skeptic and the dogmatist, *Noûs*, **34** (4), pp. 517–549.