

## **Dissolving the hard problem of consciousness**

### **Introduction**

In this paper we attempt to dissolve worries around the hard problem of consciousness by showing that there is no good argument for the existence of such a problem. The arguments for the existence of a hard problem, as defined by Chalmers (2002), come from some classic thought experiments. We are asked to imagine a scenario radically different from our experience of the world and draw the conclusion that the intrinsic qualitative nature of a mental state is independent of the structure and function of that state. The conclusion depends on the truth of identifiable key intuitions. We suggest that these intuitions are not theory neutral. Indeed not everyone shares the key intuitions. At this point there becomes a real danger of all argument collapsing to table thumping over what intuition is true. As this situation would be downright anti-philosophy we ought to ask, what would have to be the case for the key intuition to be true. In each case we suggest that the answer is that consciousness must be independent of the structure and function of mental states, i.e. that there is a hard problem. These thought experiments thus cannot serve as evidence for a hard problem; that would be question begging. At best these they show that some philosophers tend to start their approach to consciousness studies assuming that there is a hard problem.

### **What is the hard problem of consciousness?**

The main advocate of the existence of a uniquely hard problem of consciousness has been David Chalmers, and so a natural place to begin is with his characterisation of the problem. For Chalmers (2002; 1995) the hard problem is the problem of subjective experience. Specifically it is the problem of explaining the what-it-is-likeness or phenomenal properties of mental states. He asks: "how and why do physical processes give rise to experience?" (Chalmers 2002 p. 3). For Chalmers the hardness of the hard problem stems from the claim that even once we have explained all there is about the structure (e.g. composition, arrangement) and function (functional role) of a particular mental state, it seems that we will not have an explanation of the phenomenal properties of that state. No matter how much we can say about the structure and function of the state we will not know why it *seems* a certain way to be in that state, or even why it seems like anything at all.

To understand this he asks us to compare the hard problem to what he calls the 'easy' problems of consciousness. For Chalmers the easy problems of consciousness are those that it seems can be (eventually) explained by the structure and function of mental states. These are puzzles in the Kuhn-ian (1962) sense in that there is some consensus on how to approach them.

To take one example from self-consciousness research, we face a puzzle of understanding how it is that we, normal adult, humans come to experience a sense of agency for actions

we control, but not for actions we don't control. Such an explanation of this should also account for deficits in this sense observed in delusions of alien control. One prominent account attempts to explain these facts by looking at how the sense of agency is elicited. It is claimed that the sense of agency is elicited when the actual sensory feedback due to an action is represented as matching a prediction of what feedback there will be. In the case of actions that are under normal control there is a match and so a sense of agency is elicited. For actions that are not under normal control there is no, or there is a discrepant, prediction about what feedback there will be and so no match is represented and the sense of agency is not elicited. Patients suffering from delusions of alien control have a deficiency in forming the relevant predictions, so, at times, there is no match when there should be and the patient does not experience a sense of agency for actions they in fact control (Frith et al. 2000; Carruthers 2011).

From this we can see how an easy problem of consciousness may be solved, in this case by identifying certain mental representations (e.g. of actual and predicted sensory feedback) and the computational processes which act over those representations (e.g. comparison). Chalmers claim, however, is that such explanations necessarily leave something out. In this case why the sense of agency feels the way it does. It is claimed that it is not possible to get to this simply by having a better or more complete account of the kind given. For example, the comparator account typically does not have much to say about the structure of the brain activity underlying the sense of agency (although it does say a bit about the activity underlying the comparison Blakemore et al. 2003), nor is there typically much analysis of the function of the sense of agency within the cognitive system after it is elicited that goes beyond attempts to operationalize the feeling. Even if we were to fill in these gaps, we would not, it is claimed, have explained why the sense of agency feels like it does. This then constitutes the hard problem of consciousness. The methods used to make progress on the easy problems of conscious seem incapable of capturing let alone explaining the phenomenal qualities of experience.

In the following we will argue that there is no reason to believe that there is a hard problem independent of the easy problems of consciousness. When we examine the arguments for a hard problem, we see that they rely on key intuitions that are true if there is a hard problem, but not true otherwise. The only reason that can be given to believe the key intuitions is that consciousness is independent of the structure and function of mental states. As such they do not support the existence of a hard problem; to use them as such would be question begging.

### **Evaluating the evidence for the hard problem**

It is of course controversial to claim that there is a hard problem of consciousness that is distinct from the easy problems and various authors have proposed ways in which the phenomenal character of mental states may be explained by the function (e.g. Dennett

1991) or structure (Clark 1993e.g. ) of mental states. As such it is far from obvious that there is a hard problem in Chalmers sense. Although such accounts have considerable merit, we will not advocate any such account here. The reason for this is that in arguing for the existence of a hard problem, Chalmers does not attempt to show how these accounts fail to explain consciousness. Rather he attempts to show that in principle any account referring to the structure and function of mental states must fail. It is these arguments which we will address here.

These arguments all take the form of thought experiments asking us to imagine a scenario which is either impossible (Mary) or at least incredibly far from how we would normally understand the world (Zombies, Bats). The bizarreness of the scenarios alone gives us reason to question the validity of our intuitive understanding of the consequences of the scenarios. We simply have no opportunity to find out if our intuitions are true. Furthermore, as we will see below, not everyone shares the key intuitions. Indeed one of the present authors (E.S) has the intuition supporting the existence of a hard problem in some cases (Mary) but not others (Zombies). Here we will ask; “what reason can be given to believe the key intuitions are true?” Difficulties in answering this question pose problems for the advocate of a hard problem. These thought experiments have been used for a number of purposes beyond establishing the existence of a hard problem. They are used to challenge materialist understandings of the mind or to establish particular forms of dualism. We will not consider these issues here. Rather we are concerned only with the existence of a distinctly hard problem. Nor will we consider every thought experiment in the vicinity of those we consider, as many of these, such as inverted spectrum, are concerned with establishing only that the function of mental states is insufficient to account for phenomenal properties. The hard problem, in contrast concerns the stronger claim that function *along with structure* of the mental state are insufficient to account phenomenal properties.

### *Zombies*

The first thought experiment we will consider in favour of there being a hard problem involves the existence of zombies and is sometimes called the conceivability argument (Chalmers 2002). This arguments starts by asking you to imagine a zombie copy of yourself. This kind of zombie is exactly like you except that it has no experiences. The resemblance between you and the zombie must be profound. It is an atom for atom copy and does everything as you do it. We can, for example, take the zombie and attempt to induce the rubber hand illusion and it will give precisely the same reports as you, score the questionnaires the same and, indeed, a neutral observer could not tell whether you or your zombie were reporting the illusion. The zombie would seem to us to be undergoing the illusion, but it would not experience the illusion and it would have no point of view from which it could experience the illusion. Indeed the zombie would *believe* it experiences the illusion, but it would not, really, do so. Chalmers assumes that such things are conceivable

and argues from their conceivability to their existence in some possible world, i.e. that they are metaphysically possible. But, he suggests, if such a zombie were metaphysically possible (i.e. they could exist, or they exist in some possible world) then the structure and function of mental states, or indeed whole bodies, cannot explain phenomenal character. This is because there could be a being atom for atom identical to you in structure and showing exactly the same functions but with no phenomenal properties.

Here we would like to point to one key premise in this argument, namely the premise that zombies are metaphysically possible. It should be clear that the above establishes the existence of a hard problem of consciousness only if it is true that zombies are metaphysically possible. Call the intuition that zombies are metaphysically possible the key intuition of the argument. Chalmers of course spends much time, over several publications, attempting to show that the metaphysical possibility of zombies follows from their conceivability. If we grant this inference then we can consider the truth of the key intuition by considering the truth of the intuition that zombies are conceivable. If we do not grant this inference then we can examine the key intuition directly. We will begin first with the intuition that they are conceivable and then move onto their metaphysical possibility.

Neither of the present authors share the intuitions that zombies are conceivable or that they are metaphysically possible. Using Chalmers and Bourget's (2010) survey results it appears that only around 23% of philosophers are inclined to find zombies metaphysically possible and as many as 19% report that they find zombies inconceivable. These statistics do not tell us whether or not Zombies are possible, of course, but they do show that the advocate of the hard problem cannot assume the key intuitions are obviously true. In an attempt to move beyond simply butting heads on the intuition the advocate of the hard problem could argue that we ought to share their intuition. Let us examine how such an argument could work.

To start, what is it to conceive of a zombie? To some this may simply mean that thinking of a zombie does not lead to an immediately identifiable contradiction. This is fine but trivial; it is also conceivable in this sense that when I drop a cup it will hover in place or that my sister's dog will call me for a chat. Conceivability in this sense tells us precisely nothing about the world or consciousness. To attempt to find a non-trivial notion of conceivability that is argumentatively interesting let us consider a more "cognitive"<sup>1</sup> notion of conceivability. To a first approximation to conceive of a zombie is to form a mental representation of a zombie. This is obviously insufficient, however, as we can form mental representations of things which are not just impossible, but inconceivable (in any sense). For example, by saying the words "that's a square circle" to myself. So conceiving of a zombie cannot be simply a matter of reading and understanding or reciting a description of a zombie. What extra condition must be met for the conceivability of zombies to be argumentatively

---

<sup>1</sup> Thanks to Peter Menzies for the term.

interesting? We submit that we must mentally represent them in a way that coheres with our knowledge of (at least) other a priori truths. For example, saying (internally) “that is a square circle” doesn’t count as conceiving of a square circle because the notion of a square circle is incoherent with what we know of squares and what we know of circles. So, then, it seems to conceive of a zombie is to mentally represent it in a way that is coherent with what else we know to be true.<sup>2</sup>

When we have the intuition that we can conceive of a zombie, we have the intuition that we can mentally represent a zombie in a manner which is coherent with how we know the world to be. So take someone who holds that dualism is to be avoided at all costs because it leads to epiphenomenalism and thus they hold that the structure and function of mental states must be sufficient to explain phenomenal character. Such a person could not conceive of a zombie in the argumentatively interesting way because for them an atom for atom duplicate that does everything you do would necessarily have the same conscious experience as you. We see then that the intuition that zombies are conceivable is not theory neutral.

What reason then can be given to support the claim that zombies are conceivable? From the above it seems like the way to make such an argument would be to show that zombies can be imagined in a way that coheres with our broader theoretical commitments. What, then, are the theoretical commitments that would allow one to conceive of a zombie? Straightforwardly one can coherently conceive of a zombie if one believes that in principle the structure and function of mental states are insufficient to explain consciousness. As such the intuition that zombies are conceivable can only be true if one already believes that structure and function of mental states are insufficient to explain consciousness, that is, if one already believes there is a hard problem of consciousness. As such the premise that zombies are conceivable cannot appear in an argument for the existence of a hard problem, to include it is to beg the question. Given this and given that we granted Chalmers the claim that metaphysical possibility follows from conceivability it seems the argument from conceivability for a hard problem of consciousness does not go through. One cannot take this argument and thought experiment to show that there is a hard problem of

---

<sup>2</sup> This cognitive notion of conceivability seems to raise two major problems: first of all it seems to make it impossible to conceive of fictitious entities like vampires; and secondly it seems to be inappropriately cognitive and at odds with the more “metaphysical” notion of conceivability that is relevant to this debate.

A possible objection here is that it appears that such a notion of conceivability would make things like vampires inconceivable. This would be because we know that they don’t exist and that something cannot be dead and animate at the same time. So to think of them does not cohere with what else we know. To make this objection, however, would be to fail to notice that suspending such beliefs and the notion of a vampire both cohere perfectly well with our notion of fiction or pretence. On the idea of conceivability which we are using here vampires are conceivable as fictional entities.

Secondly one could complain that what we are interested in here is not our beliefs about consciousness, but rather the a priori truths about consciousness. We are not interested in denying this claim, rather our focus on beliefs is a means to highlight the fact that, given our uncertain epistemic situation, justification of these beliefs is necessary for them to count as knowledge.

consciousness, unless one assumes the existence of a hard problem at the first premise- that zombies are conceivable.

Suppose though that we were to drop the first premise, that zombies are conceivable, and just work from the key intuition that zombies are the kind of thing which could exist; that they are metaphysically possible. What reason can be given to believe that zombies are metaphysically possible? In this case we can ask, what does the world have to be like such that this intuition is true? Zombies cannot exist if the structure and function of mental states are necessary and sufficient to explain consciousness, as in that case, in every possible world, something that was an atom for atom duplicate of you doing all of the same things as you would have the same experiences as you. We see the truth of the key intuition is not theory independent.

Again, we should ask: what must the world or a possible world be like in order for the key intuition to be true- i.e. for zombies to be metaphysically possible? The answer, again, is clear enough. The structure and function of mental states (indeed whole beings) must be insufficient to explain phenomenal properties. In other words, for the key intuition in the thought experiment to be true, and it must be true to support the conclusion of Chalmers argument, it must be the case that there is a hard problem of consciousness. The truth of intuition is thereby dependent on there being a hard problem and cannot be justified without assuming a hard problem. As such this intuition ought not feature in any argument for the existence of a hard problem. To include it is to beg the question.

### *Mary*

The thought experiment about Mary, the colour blind super scientist, and the accompanying knowledge argument were originally introduced to argue against materialism and for a kind of property dualism (Jackson 1982). The argument can however, also be used as an argument for the existence of a hard problem of consciousness (Chalmers 2002). The core of the knowledge argument is to claim that there are some facts about consciousness which are not knowable, deducible or explainable by facts about the structure and function of mental states, brains and indeed the organism and its environment. If true then there is a hard problem of consciousness. In an attempt to establish this we are asked to imagine a super scientist, Mary, who is raised in an entirely black and white environment and so never has any colour experiences. Mary exists at some hypothetical time in the future when scientists and philosophers have discovered and formed correct theories about colour vision. Everything is known about how light interacts with objects and eyes, how visual discriminations are made, how one visually imagines objects et cetera, et cetera. She has perfect knowledge of the structure and function of visual states, and indeed anything else relevant to the formation of those states. Mary, in virtue of being a super scientist, can possess all of this knowledge, that would in fact be distributed across several identifiable scientific communities. In virtue of being a perfect reasoner she can deduce, or at least

explain everything that follows from the knowledge she acquires. What then, we are asked, will happen when she is let out into the coloured world and is greeted by, say, a scarlet macaw? Will she learn something that she did not know before, namely what it is like to see red? The intuition which we are asked to share is that yes, she will learn something new. If true then she has not learned all there is about colour visual experiences despite her perfect knowledge of the structure and function of visual mental states. Thus, the structure and function of such states is insufficient to know, deduce or perhaps even explain<sup>3</sup> the phenomenal character of such states. Thus, it is concluded, there is a hard problem of consciousness.

The key intuition in accepting this argument is the intuition that Mary learns something new upon having colour experiences. Of course there are any number of reasons, mostly to do with the finite and plastic nature of the brain, to believe that such a scenario could never actually occur. But, to the advocate of the hard problem this is a trivial response. So let us take the intuition seriously. That there is a hard problem of consciousness can only be established by this argument if it is true that Mary would learn something new upon having colour experiences. Like the intuition that zombies are metaphysically possible, this intuition is not universally held or obviously true. One of us (E.S.) has the strong intuition that Mary learns something new; the other (G.C.) regularly equivocates and so doesn't trust his intuitions about the scenario. More prominently Dennett (1991) appears not to share the key intuition. He advocates an account on which certain *functions* of mental states constitute consciousness of that state. On Dennett's account on knowing this functional account of consciousness Mary knows what it is like to experience colours. As he puts it so vividly, if you think Mary learns something new: "you are simply not following the directions" of the thought experiment (Dennett 1991 399). We see again that as well as not being universally held, the truth of the intuition that Mary learns a new fact is not theory neutral.

What reason can be given to show that Mary learns something new when she leaves her black and white environment? The advocate of the hard problem could attempt to convince those who don't have the intuition that they should by asking what must the world be like in order for Mary to learn something new when she enters the coloured world, and then attempting to show that the world is, in fact, like that. Such a move, however, would fail at the first step. For Mary to learn something new it must be the case that all of her knowledge of the structure and function of visual states is insufficient to provide knowledge of, allow her to deduce and perhaps even explain the phenomenal character of coloured experiences. If her knowledge were sufficient for this she would not learn something new. In other words the key intuition can only be true assuming there is a hard problem of consciousness. As

---

<sup>3</sup> However, see Schier (2007) for an account on which there are physical facts, knowable in only one way (and thus not known by someone, like Mary, who has no access to a particular way of knowing) and yet still explainable.

such this intuition cannot enter as a premise in the argument for the existence of a hard problem. As the knowledge argument depends on the truth of this intuition to establish the existence of a hard problem, the knowledge argument fails to establish the existence of a hard problem as it assumes it at the premise of the key intuition.

### *Bats*

The problems that Nagel (1974) has raised are many and varied. Some may be seen as forerunners of the zombie and knowledge arguments so we will not consider these further here. There is, however, an additional problem which we may also consider as evidence for there being a hard problem of consciousness. This problem has to do with the appearance and reality of consciousness. Nagel (1974 448) asks us “Does it make sense...to ask what my experience is *really* like, as opposed to how they appear to me?”. With this question Nagel communicates a reasonable intuition that consciousness is composed entirely by appearances to the subject. For Nagel this creates problems for materialism, but we can also use similar reasoning as evidence for the hard problem. If the intuition that consciousness is composed entirely of appearances to a subject is true then this would seem to pose two reasons to believe in a hard problem.

First, in doing science, say by discovering the structure and function of mental states, we leave behind their subjective appearances in an attempt to discover objective reality. To use one of Nagel’s examples when we say what lightning is we are less interested in how it appears to us, then the reality of electrical discharge and the compression and rarefaction of air (Nagel 1974 443). We leave behind how lightning appears to us in order to capture what it really is. However, if we leave behind the appearance of consciousness, say in order to study the structure and function of mental states, we leave behind the entire phenomenon of consciousness. Thus, the structure and function of mental states are independent of what it is like to undergo those states, and thus there is a hard problem of consciousness.

The second way in which consciousness being entirely composed of appearances to the subject leads to the hard problem is when we consider conscious experiences of a different type than our own. To illustrate this Nagel chose the vivid example of bats as they have a sonar system which presumably involves experiences of a radically different type to what any of us can have (Nagel 1974 438, 441). Again, for Nagel if we study a bat we can learn a lot about it, especially the structure and function of its nervous system and mental states (cf. Akins 1993). But we cannot learn what it is like for the bat to be in those states. The reason for this is that consciousness of those states is wholly constituted by the appearance of those states to the bat. This is only accessible from the bat’s point of view. When we study the bat we can do so only objectively, say by studying the structure and function of the bat’s nervous system. But in doing so we necessarily leave behind the appearances to the bat. Thus, for the advocate of the hard problem, in studying the structure and function



of the bat's nervous system and its mental states we leave behind what is constitutive of consciousness, namely the appearance of those states to the bat.

The key intuition that we wish to draw attention to here is that consciousness is wholly constituted by appearances to the subject. The truth of this intuition, like the other key intuitions is theory dependent. Nagel himself recognises the theory dependence of the key intuition when he states that a conceptual change would allow us a way out of the problems he raises (Nagel 1974 447). Nagel himself, of course, asserted that problems like the hard problem may not be permanent, in principle problems, but rather problems of how we conceptualise consciousness (Nagel 1974 446). We can see the theory dependence of the truth of this intuition when we ask, what would the world have to be like for this intuition to be true? It would not be true if the structure and function of mental states were sufficient to explain consciousness. For in that case, there is more to consciousness than appearance to the subject, namely the structure and function of mental states. Indeed it is not true even if the structure and function of mental states are only *a part* of the explanation of consciousness, nor if there is merely a neural underpinning of consciousness; as in each case there is more to consciousness than appearance to the subject.

For this intuition to be true, then, it must be the case that consciousness is independent of the structure and function of mental states. In other words the truth of this intuition depends on an acceptance of there being a hard problem of consciousness. As such neither of the arguments presented in this section can be used to justify the existence of a hard problem of consciousness, both depend on their being a hard problem to establish the truth of the key intuition.<sup>4</sup>

### **Dissolving the hard problem**

We have seen then that each of the thought experiments thought to establish a hard problem of consciousness depend on the truth of identifiable key intuitions. When we ask, however, under what conditions these intuitions could be true, we see that they are true if the structure and function of mental states does not suffice to explain or provide knowledge of phenomenal properties. That is, the key intuitions are true if there is a hard problem of consciousness and false if phenomenal properties are to be explained by the structure and function of mental states. The presuppositions of the arguments for the existence of a hard problem are therefore not theory independent and, indeed, are not neutral to the existence

---

<sup>4</sup> This intuition is also a part for a less formal argument for the existence of a hard problem. It may be claimed that, based on a mere examination of one's own experiences, that it is just obvious that consciousness poses a different kind of problem to other mental phenomena, because there is a gap (either explanatory or metaphysical) between what it is like to have an experience and the material world. However, this is only obvious if one believes that in examining one's own experience one has correct and complete knowledge of one's consciousness, otherwise there could be any number of things which fill the gap. In turn this is true only if everything about consciousness is exhausted by appearances to the subject. As we have seen the truth of this intuition is conditional on the existence of a hard problem and so this less formal argument cannot be used as evidence for a hard problem.

of the hard problem. The truth of the key intuitions presuppose the existence of a hard problem and thus cannot be used to establish its existence.

At best, if one is convinced by the thought experiments that there is a hard problem all that can be established by this is that one *already* held intuitions which are consistent with there being a hard problem. This is, of course, no use in establishing the truth of the intuitions, which we have seen are not universally shared.

Additionally, as we have seen, one's evaluation of the truth of the key intuitions can change depending on one's theoretical commitments. In particular we have seen that if one held a theory on which structure and function were sufficient (in principle) to explain phenomenal properties then the key intuitions are false. Our acceptance of the conclusions of the thought experiments is therefore also theory dependent. So instead of letting our conclusions on the thought experiments guide our theories of consciousness, we should let our theories of consciousness guide our conclusions from the thought experiments. Does Mary learn something new upon leaving her black and white room? Could Zombies really exist? What is it like to be a bat? The answer to all of these questions is; we do not know until there is a generally accepted theory of consciousness. This is an exciting and difficult project, but there is no reason to believe it is 'hard' in Chalmers' sense.

## References

- Akins, K. (1993). What is it like to be boring and myopic? . Dennett and his Critics. B. Dahlbom, Blackwell.
- Blakemore, S.-J., D. A. Oakly, et al. (2003). "Delusions of Alien Control in the Normal Brain." Neuropsychologia **41**: 1058-1067.
- Carruthers, G. (2011). "The case for the comparator model as an explanation of the sense of agency and its breakdowns." Consciousness and Cognition.
- Chalmers, D. J. (1995). "Facing up to the problem of consciousness." Journal of Consciousness Studies **2**(3): 200-219.
- Chalmers, D. J. (2002). Consciousness and Its Place in Nature. Philosophy of Mind: Classic and Contemporary Readings. D. J. Chalmers, Oxford.
- Chalmers, D. J. and D. Bourget. (2010). "The philpapers survey." from [http://philpapers.org/surveys/results.pl?affil=Target+faculty&areas0=0&areas\\_max=1&grain=fine](http://philpapers.org/surveys/results.pl?affil=Target+faculty&areas0=0&areas_max=1&grain=fine).
- Clark, A. (1993). Sensory Qualities, Clarendon Library of Logic and Philosophy.
- Dennett, D. C. (1991). Consciousness Explained. New York, Penguin Books.
- Frith, C. D., S.-J. Blakemore, et al. (2000). "Abnormalities in the Awareness and Control of Action." Philosophical Transactions of the Royal Society of London Series B - Biological Sciences **355**: 1771-1788.
- Jackson, F. (1982). "Epiphenomenal Qualia." Philosophical Quarterly **32**: 127-136.
- Kuhn, T. (1962). the Structure of Scientific Revolutions, University of Chicago Press.
- Nagel, T. (1974). "What is it Like to be a Bat?" Philosophical Review **83**: 435-450.
- Schier, E. (2007). Emergence, colour and the knowledge argument: so what if Mary didn't know. PhD, University of Adelaide. <http://hdl.handle.net/2440/57408>.

