Phenomenal Consciousness Must Be Sharp

Joshua O’Rourke

Introduction

Is a mosquito phenomenally conscious? This is a difficult question, and it is not clear how we would go about settling it, but one thing it seems we can say, even in our present state of knowledge, is that the answer is determinately yes or no. The mosquito might possess only a dim and simple sort of consciousness, but to be dimly and simply conscious is still to be conscious. Intuitively, there are no borderline cases of consciousness. This means that a fully correct theory cannot use a vague concept to pick out the property of being conscious.\(^1\) Most physicalist theories of consciousness fail to satisfy this condition, however, because they use inexact concepts, drawn either from folk discourse or from the non-fundamental sciences, to pick out the physical phenomenon that is supposed to be identical to consciousness.

These considerations give us the following argument schema:

(Premise 1) There are no borderline cases of consciousness.

\(^1\)I assume that a concept is vague if and only if it has borderline cases.
(Premise 2) Theory $T$ picks out consciousness in vague terms.

(Conclusion) $T$ is not the correct theory of consciousness.

A fully worked-out theory of consciousness will give us a predicate, $P$, such that a being $x$ is conscious if and only if $Px$.\(^2\) Given that consciousness does not have any borderline cases, this can only work as an account of consciousness if $P$ is likewise non-vague. Of course, no physicalist has given a theory detailed enough to say exactly what $P$ would be, but we can say in general terms how $P$ would be determined on different proposals.

It is worth taking a moment here to clarify how I am thinking about vagueness. Theories of vagueness fall into three major categories, semantic indecision theories, epistemic theories, and ontic theories. On a semantic indecision theory of vagueness,\(^3\) a vague concept or predicate does not succeed in picking out a single property because our dispositions to apply it are not exact enough to decide between a family of very similar properties. On an epistemic theory of vagueness,\(^4\) vagueness is a special kind of ignorance. Here, again, we have a family of very similar properties, but the epistemist thinks that the metasemantic rules are sensitive enough to ensure that a vague concept or predicate nonetheless singles out just one of those prop-

\(^2\) Depending on the ambition of our theory, we might want more from this biconditional than just its truth. Below, I distinguish between theories that seek to analyze our concept of consciousness and theories that merely seek to tell us which phenomenon is picked out by our concept of consciousness.

\(^3\) Fine (1975) and Lewis (1986) are two of the most influential defenders of this sort of account, but it is common enough to count as the current orthodoxy.

\(^4\) The epistemic theory of vagueness originated with Williamson (1994), and he remains the philosopher most associated with it.
erties. However, because the properties in question are all so similar, the concept or predicate could easily have picked out a different one of those properties. Thus, it is impossible for us to know exactly which one of those properties our concept or predicate picks out.\(^5\) On an ontic theory of vagueness,\(^6\) a concept or predicate gets to be vague in virtue of picking out a single property that is itself vague.\(^7\) Ontic theories locate vagueness in the world, while semantic indecision and epistemic theories locate vagueness in our representations of the world or in our knowledge about those representations. I think the argument schema given above applies no matter which theory of vagueness is correct, so, for the purposes of this paper, I will not take a stand on this issue.

Are instances of this argument schema valid? First, consider instances on which theory \(T\) purports to analyze our concept of consciousness. This kind of theory is supposed to tell us what we meant by “consciousness” all along. Thus, an adequacy condition on such a theory is that \(Px\) and “\(x\) is conscious” mean the same thing, but a vague predicate and a non-vague predicate clearly differ in meaning. This shows that, when applied to an analytic theory, the schema produces a valid argument. Next, consider

\(^5\)This has to do with Williamson’s link between knowledge and safety. The idea is that, although we could form a true belief about which property our concept or predicate picks out, there would always be very nearby possible worlds in which this belief is false. This means that such a belief is not safe and therefore does not amount to knowledge.

\(^6\)Defenders of ontic theories of vagueness, or at least the possibility of such theories, include van Inwagen (1990), Rosen and Smith (2004), Williams (2008), Barnes (Barnes 2010; Barnes 2014), Barnes and Williams (2011), and Wilson (2013).

\(^7\)Of course, one could mix and match these theories. Perhaps different accounts are needed for different varieties of vague phenomena.
instances on which \( T \) merely purports to identify the property that is picked out by our concept of consciousness. These instances are also valid, but the reasons why depend on which theory of vagueness is true. On an ontic theory of vagueness, the situation is very straightforward. A vague predicate picks out a vague property, but the property of being conscious is not vague, so \( P \) cannot be vague. On a semantic indecision theory, the situation is also fairly straightforward. If our concept of consciousness is not vague, then it picks out a single property. If \( P \) is vague, then it fails to pick out a single property. Thus, a vague predicate cannot tell us which property is picked out by our concept of consciousness. On an epistemic theory of vagueness, things are a bit more complicated because our concept of consciousness could be non-vague, \( P \) could be vague, and yet \( P \) really could pick out the same property as our concept of consciousness.\(^8\) This is a fair point, but note that a theory like this cannot tell us what property is picked out by our concept of consciousness because, given the truth of epistemicism, it is impossible for us to know what property is picked out by \( P \). However, because our concept of consciousness is not vague, it should be possible to know what property it picks out. I claim, then, that the correct theory of consciousness will pick out consciousness in non-vague terms, even if epistemicism is true.\(^9\)

\(^8\)The biconditional “\( x \) is conscious if and only if \( Px \)” would still have borderline cases, though, so its universal closure would not be \textit{determinately} true.

\(^9\)What about a physicalist who thinks we just lack the conceptual tools to drill down to a perfectly precise theory of consciousness? This sort of physicalist could agree that consciousness has no borderline cases while disagreeing that the best theory of consciousness must dispense with vagueness, if by “best theory” we mean “best theory available to creatures like us.” Such a view is possible, of course, but it is not very satisfying. Fur-
One worry here is that our concept of consciousness might be like the pre-modern concept of life. People used to assume, or so the story goes, that there must be sharp cutoffs between the living and the non-living, but, over time, they arrived at a more scientifically informed concept of life that identifies being alive with having a vague cluster of vaguely defined capacities.\textsuperscript{10} The concept of life in use at the beginning of this story is defective at least in part because of its sharpness. There are deep questions about how exactly to make this point,\textsuperscript{11} but I will mostly try to skate over them here. Whatever the correct theory of concepts is, it must be capable of capturing the following point. Life is in fact a vague phenomenon, so the correct theory of life will be given in vague terms, however people originally conceived of it. One might argue that consciousness, like life, is a vague phenomenon, so if our concept of consciousness is sharp, then it is defective. In this paper, I will argue that not only is our concept of consciousness sharp, like the pre-modern concept

\textsuperscript{10}On one way of telling this story, one concept persisted through a significant change in its application conditions. On another way of telling this story, one concept was replaced by a different concept. I think we need to build some tolerance into the application conditions of our concepts, otherwise it would be hard to resist the conclusion that we are all using conceptual schemes that are totally incommensurable with one another, but it will not be important for my purposes in this paper to take a stand on which way of telling the story is correct.

\textsuperscript{11}Did the sharp concept of life apply to things, like people and dogs, that its users were unreservedly disposed to apply it to, or was its actual extension empty because vitalism is false? A distinct but related question is whether there is a clear line between a concept’s application conditions and the assumptions about the world that are implicit in our use of it.
of life, it is correct in being so. Consciousness, as it exists in the world, is sharp.

In this paper, I argue that the argument schema presented above can be applied to the major contenders among physicalist theories of consciousness. In section 1, I discuss the case for the first premise. In section 2, I argue that the second premise is true of two broad classes of physicalist theories, those that pick out consciousness in functional terms and those that pick out consciousness in neuro-physiological terms. In section 3, I canvass the theories that this argument does not rule out.

Many philosophers have acknowledged both the force of the intuition that consciousness cannot be vague and the fact that this is hard to square with a physicalist worldview on which macroscopic objects are just clouds of particles.\textsuperscript{12} Recently, Antony (2006a; 2008) and Simon (2017) have argued that our concept of consciousness is sharp, although they have not addressed whether it is correct in being so. I discuss these arguments in section 1. Antony (2006b) has also argued that if our concept of consciousness is sharp and correct in being so, then many leading theories of consciousness must be false. I discuss these arguments in section 2.

1 Consciousness Cannot Be Vague

I begin this section by arguing that our present concept of consciousness is sharp. One might grant this while maintaining that consciousness itself, the phenomenon we are trying to capture by our use of the concept, must be vague, so I go on to give a pair of related arguments for the conclusion that consciousness itself is sharp. I close the section by discussing the source of our intuition that premise 1 must be correct.

1.1 Our Concept of Consciousness Is Not Vague

Many of our concepts have borderline cases. This includes the standard examples, like TALL and BALD, but it also includes concepts we do not normally think of as vague, like CHAIR.\(^{13}\) Suppose someone is assembling a chair but loses interest part of the way through. Is the resulting object a chair or not? We can easily imagine borderline cases, i.e., cases in which it is indeterminate whether an object possesses enough of the features typical of chairs to be counted as a chair. What if it only has three legs and cannot stand on its own? What if the back is not attached? We could legislate an answer in these sorts of cases, but the point is that the concept we actually use does not force us one way or the other.\(^ {14}\)

---

\(^{13}\) When referring to concepts, I either use the “concept of” locution or small caps, as I do here.

\(^{14}\) Following Alston (1964), Antony (2008) distinguishes between degree and combinatory vagueness. Degree vagueness arises for concepts like TALL that involve variation along a single dimension when it is not clear where to draw the line. Combinatory vagueness arises for cluster concepts like CHAIR when it is not clear exactly which features have to
The concept CONSCIOUS is not like CHAIR in this way. Following Nagel (1974), I claim that there is just one feature that determines whether an object is phenomenally conscious: there being something it feels like to be that object, and this feature is always either completely present or completely absent. For a similar sort of case in the physical realm, consider the concept of having mass. Something that has only a very small mass, such as a neutrino, nonetheless has mass without qualification. Likewise, an object that has any degree of phenomenal experience is conscious full stop. There is no such thing as being sort of conscious.

In an effort to convince us that a borderline case of consciousness must exist somewhere, an objector might try to construct a sorites series for consciousness, i.e., a sequence of cases starting with a clear case of consciousness and ending with a clear case of non-consciousness such that the difference between adjacent cases seems insufficient to ground a change from consciousness to non-consciousness. This could be done either from the outside, in physical or functional terms, or from the inside, in phenomenal terms. For an example of the first sort of series, consider a sequence of living things with simpler

---

15 What I say here is not strictly correct because there could be borderline cases of MASSIVE in possible worlds that have slightly different physical laws than our own. If, in such a world, there is a quantity that behaves kind of like mass behaves in our world, it might be indeterminate whether something that possesses that quantity should fall under our concept MASSIVE. Although I do not think that consciousness can exhibit this sort of vagueness either, I wish to bracket the issue here. The analogy between CONSCIOUS and MASSIVE goes through if we restrict ourselves to nomologically possible worlds.
and simpler information-processing capabilities. Perhaps this sequence starts with humans and ends with bacteria. We must reach a point, it seems, after which all successive beings are just too simple to sustain consciousness, but it is hard to know where that point would be. For an example of the second sort of series, consider the process of being put under anesthesia. At what point in this process do you pass from being determinately conscious to being determinately non-conscious? Again, it is hard to say.

Let us first consider series described in physical or functional terms. Antony (2006a) points out that the sort of indecision we display toward these kinds of difficult cases is very different from the sort of indecision that we display in paradigm cases of vagueness. When I am not sure whether I should call a person tall, I do not feel that this is due to my ignorance. After all, I might know the person’s exact height. In a paradigm case of vagueness, there is no pull to think that there must be a determinate answer. However, when I am not sure whether I should call a mosquito conscious, this seems to be due to a lack of relevant information. Is there something it feels like to be a mosquito? If so, then the mosquito is conscious. If not, then it is not. Our reactions toward cases where it is difficult to say whether a being described from the outside is conscious are very different from our reactions toward typical borderline cases of vague predicates. Antony (2008) and Simon (2017) go on to argue that if a competent user of a concept is not disposed to react to difficult cases in ways that are characteristic of vague
concepts, then that concept cannot be vague.\footnote{Antony and Simon differ on what ways of using a concept are characteristic of vagueness, but these differences are not relevant here.} Their idea is that a competent user must be sensitive to the fact that a concept is vague.\footnote{One could think that the explanatory relation runs in either direction here. Perhaps a concept counts as vague because competent users have the right sorts of dispositions toward it, or perhaps a user counts as competent in part because she responds to vague concepts in the right way. In either case, the argument goes through.} Because we are competent with CONSCIOUS, the fact that we do not respond to difficult cases in ways characteristic of vagueness shows that CONSCIOUS cannot be vague.

Is it really true that a concept cannot be vague without people knowing it? Think again of the pre-modern concept of life. One could argue that this concept was vague even before people knew that fact. The rules for reference-fixing presumably work by maximizing the fit between our usage of our concepts and the actual structure of the world. Thus, the way we are disposed to use concepts does not simply dictate their content. The world has a part to play as well. Perhaps the pre-modern concept of life succeeded in picking out the actual phenomenon of life, even though people’s theoretical beliefs about life were largely false. If this is right, then because the actual phenomenon of life is vague, we should think that the concept of life was vague even before people came to realize this. Antony and Simon could argue that these benighted people were not competent with the concept of life, but, of course, they thought they were. If we were incompetent with CONSCIOUS for similar reasons, we would not know it. Note here that neither
Antony nor Simon takes themselves to have argued for the conclusion that consciousness, the worldly phenomenon, is sharp. They think we could be mistaken about this. I do not know whether this point is best put by saying that our present concept of consciousness is sharp but could be defective in being so or by saying that we could be mistaken in assuming that our concept is sharp, but I think the general point is clear.

In the next subsection, I set myself the task of arguing that we are not mistaken in this way about the nature of consciousness. I will, however, make one point here on behalf of the way Antony and Simon are thinking about things. If the pre-modern concept of life was vague, then it should be possible to describe a borderline case in enough detail to convince a user of that concept that it would be a borderline case if it existed. To preserve the conviction that there are no borderline cases of life, then, this person would have to believe that the described case is impossible. This is a disanalogy with the discussion in the previous paragraph. No one is questioning whether it is possible for mosquitoes to exist. We are simply unwilling to treat them as if they are borderline cases of consciousness.

Moving on now to series described in phenomenal terms, Antony and Simon argue that we cannot describe a borderline case of consciousness from the inside. This is because any attempt to clearly describe such a case will have to answer the question of whether there is something it feels like to be the subject. Consider again being put under anesthetic. At each moment, you will either feel something or feel nothing. If you feel anything at all,
then you are still conscious. If you feel nothing, then you are no longer conscious. Of course, anesthetic erodes the functional capacities of your brain as well as its experiential capacities, so it might be impossible for you to say, either at the time or when you look back on it later, just when your experience winks out, but this does not mean that you undergo anything like a borderline conscious state. If you could examine each experience with your discriminative capacities intact, perhaps by having a distinct memory of it, then you would be able to see that it either feels like something or it does not.

There is a worry here if one agrees with Goodman (1951) that two different experiences could be phenomenally indiscriminable. Perhaps, even under ideal conditions, we might not be able to tell a state that was just barely conscious from a state that had no phenomenal character at all. How, then, could we know that there are no borderline cases between the two? The issue with this line of thinking is that phenomenal differences have to be differences in how things feel. Thus, the difference between being phenomenally conscious and not has to be a felt difference. Whether we can detect the difference, where detection involves high-level cognitive capacities, is irrelevant.

1.2 Consciousness in the World Is Not Vague

I think it is hard to deny that our present concept CONSCIOUS is sharp, or at least that our implicit theory of consciousness implies that it has no borderline cases, but why should we think that consciousness itself is sharp?
Antony and Simon admit that, for all they have said, CONSCIOUS might turn out to be like the pre-modern concept of life. The fact that we are presently disposed to assume that there are no borderline cases of consciousness gives us only a weak reason for thinking that consciousness itself, the phenomenon we are trying to capture with our concept, is sharp. To offer a stronger reason for this belief, then, I will articulate a pair of related arguments (or perhaps two versions of a single argument) that I find convincing. The first of these arguments likely begs the question against the physicalist, but the second might be able to gain some purchase on the physicalistically inclined.

The arguments proceed from the following sort of picture. A thing is conscious if and only if it is in some maximally specific phenomenal state.\textsuperscript{18} This much I take to be uncontroversial. The first argument runs as follows. The maximally specific phenomenal states, or some spanning subset of them if they turn out to be interdefinable, are fundamental, and a fundamental feature of the world must be sharp. Put another way, it is a basic tenet of scientific inquiry that when a phenomenon turns out to be vague, this shows us that we have not yet reached the fundamental level. If what we want is a fundamental description of the world, then we must seek a more exact vocabulary. Thus, for each maximally specific phenomenal state, it cannot be vague whether an object is in that state, so it cannot be vague whether an\textsuperscript{18} Depending on which theory of vagueness is true, assuming that a concept picks out a property might prejudice the question of whether it is vague. I thus use “state” in an effort to remain neutral on whether phenomenal concepts pick out properties or not. In my terminology, something’s being in state $F$ just means that it falls under our concept of $F$.
object is in one or another of those states. This implies, then, that it cannot be vague whether or not an object is conscious.\textsuperscript{19}

Of course, the claim that some phenomenal states are fundamental is tantamount to dualism. Instead of assuming they are fundamental, then, the second argument uses the premise that the maximally specific phenomenal states are perfectly natural. A state is perfectly natural if and only if, when two objects are both in it, they are thereby exactly similar in some respect.\textsuperscript{20} It does not seem possible, however, for a vague state to ensure exact similarity. Intuitively, this is because there is some spread in how a vague predicate describes something to be. Now, almost all predicates, even ones that are supposed to pick out perfectly natural states, leave open questions about the objects to which those predicates apply. Two objects might share the perfectly determinant shade crimson, rendering them exactly alike in color, while being totally different in other ways. Perhaps one is a fire truck and the other is a stop sign. This kind of spread in the way the predicate

\textsuperscript{19}One might object here that quantum mechanics implies that the fundamental level is in fact not sharp. What if a system is in a superposition of two states, one in which consciousness is present and the other in which it is absent? Consider Schrödinger’s cat. If we can set up a situation in which it is indeterminate whether the cat is alive or dead, then surely that would be a situation in which it is indeterminate whether the cat is conscious or not. The first thing to say here is that even if a system is in a superposition with respect to a particular observable, it is still in a perfectly determinate state, just not an eigenstate. Quantum mechanics, on one of way of thinking about things, has just given us a new set of exotic, although perfectly determinate, ways of being with respect to its observables. The second thing to say is that even if quantum superpositions are best understood in terms of indeterminacy, I could simply bracket these cases. If I am forced to say that consciousness is never vague, except in certain cases of quantum weirdness, then the main arguments of the paper would still go through.

\textsuperscript{20}This idea comes from Lewis (1983).

14
“is crimson_17” describes something to be consistent with its picking out a perfectly natural state. Contrast this with the predicate “is crimson_17 or crimson_18.” This predicate cannot pick out a perfectly natural state because two objects could share it without being exactly alike in color. A vague predicate seems more like the second case than the first because it cannot ensure that different objects falling under the predicate are exactly alike with respect to the underlying dimension of similarity. Sharing a vague state ensures vague similarity, but that is not enough. This shows that the maximally specific phenomenal states, assuming they are perfectly natural, cannot be vague. Thus, it cannot be vague whether or not an object is conscious.\textsuperscript{21}

If the argument in the previous paragraph is unconvincing, Dorr and Hawthorne (2013) offer a totally separate path to the conclusion that perfectly natural states cannot be vague. Their argument uses the premise that perfectly natural states exhibit a phenomenon called reference magnetism.\textsuperscript{22} In other words, perfectly natural states are supposed to be particularly easy to refer to. However, vague states nest within a halo of very similar properties, rendering them difficult to refer to. Vague predicates are vague precisely because we cannot drill down and say exactly what we mean by them.\textsuperscript{23} Thus, vagueness and reference magnetism seem to be at odds with one an-

\textsuperscript{21}Given a vague predicate $P$, could the restricted predicate that only applies to things that are determinately $P$ pick out a perfectly natural state? Perhaps, but only if the restricted predicate is non-vague.

\textsuperscript{22}This idea is also present in Lewis (1983).

\textsuperscript{23}It is less clear how this is supposed to apply to ontic theories of vagueness, but note that even properties that are themselves vague will be located within a family of very similar sharpenings and loosenings.
other, and we should assume that vague predicates do not refer to perfectly
natural states.

Fundamentality and perfect naturalness are both ways of trying to get
at the idea of carving nature at its joints, and some philosophers, like Zim-
merman (2010), assume that a state is perfectly natural if and only if it is
fundamental. Hence, I am not sure whether we have two distinct arguments
here or just two ways of putting the same argument. In any case, if we are
dealing with joints in nature, then the states in question must be present in
the world itself. I will tip my hand here and say that I agree with semantic
indecision theorists and epistemicists that vagueness is a matter of how we
represent the world, so a distinction that is really present in the world is
not the sort of thing that can have borderline cases. I grant that the defender
of ontic vagueness might have the tools to resist the idea that the fundamen-
tal/perfectly natural phenomenal states are non-vague. Nonetheless, nothing
about the picture I have presented is inconsistent with ontic vagueness, and I
think that those who believe in ontic vagueness should at least concede that

\[24\] Very briefly, I think this because, given any vague situation, there is always a per-
fectly precise description of that situation available, and I believe that this kind of precise
description should always be metaphysically privileged over the corresponding vague de-
scription. Consider the defense of ontic vagueness presented by Rosen and Smith (2004). They
do things in terms of degrees of truth, which makes it clear that we can replace a
vague description in terms of vague property \(P\) with a precise description in terms of the
family of precise properties \(\{x \text{ has } P \text{ to degree } y : 0 < y < 1\}\). Even if we are not presented
with a description in terms of degrees of truth, the presence of higher-order vagueness will
ensure that the different candidates for bearing \(P\) are ordered by the better-candidate-for-
bearing-\(P\) relation in such a way as to ensure that a precise degree-theoretic description is
available. Given that it is open to us to eliminate ontic vagueness, then, I think we should
do so.
our concept’s being sharp gives us some reason for thinking that the property of being conscious is also sharp.

Putting the argument that consciousness is not a vague state in terms of naturalness might make it a little easier for the physicalist to swallow. It seems hard to deny that two things that share the same maximally specific phenomenal state are exactly similar in some respect. Of course, the physicalist could hold the line by denying it all the same. I suspect that the arguments just given will not do much to convince the committed physicalist. However, they do show that the physicalist must bear a theoretical cost, either by denying that psychological similarity is a metaphysically heavy-duty dimension of similarity or by denying that perfectly natural states must be sharp.

1.3 How Can We Know This?

I close this section by discussing where our intuition that consciousness must be sharp might come from and how it might be justified. Perhaps it seems like borderline cases of consciousness are impossible only because we cannot imagine what they would be like from the first-person perspective. This would be a problem because the inability to imagine something from the first-person perspective gives us no reason for thinking that it is impossible. I cannot imagine what it is like to be a bat, but I am not thereby justified in thinking that bats cannot have conscious states. This is a fair point, but I do not think that our inability to imagine borderline cases of consciousness
from the first-person perspective could be the source of our intuition that such cases are impossible because this kind of inability does not, in general, exert any intuitive pull on us at all. When I try and fail to imagine what the conscious states of bats are like, I am not moved to say that bats cannot be conscious. Likewise, when I try and fail to imagine what a world without consciousness would be like from the inside, I feel no intuitive pressure to say that such a world is impossible. Thus, the fact that I cannot imagine borderline cases of consciousness from the first-person perspective cannot be the explanation for why I judge them to be impossible.

Setting aside first-person conceivability, perhaps we think that borderline cases of consciousness are impossible because we cannot imagine them from the third-person perspective. Again, the inability to imagine something typically says more about the limits of one’s imagination than about objective possibilities, so this would provide only very weak justification, if any at all, for the intuition that borderline cases of consciousness are impossible. However, I claim that a lack of third-person conceivability cannot be the source of this intuition because we actually can imagine borderline cases of consciousness from the third-person perspective, inasmuch as we can imagine determinate cases and non-cases of consciousness. When we imagine a situation from the third-person perspective, we imagine it from the outside, but, conceptually, we do not take anything that is observable from the outside to definitively imply consciousness.25 This means that no matter what

25 Certain philosophers of perception, like Siegel (2017), believe that we can literally see
we observe about an object, there will always be an open question about whether it is conscious. Likewise, it means that predicating consciousness of something does not entail anything about its observable properties. Thus, imagining that something is conscious from the third-person perspective is compatible with imagining anything at all about its observable properties. This means that, in order to imagine that something is conscious from the third-person perspective, all we need to do is imagine that thing and then predicate consciousness of it in our imagination. If this is all we are doing when we imagine that something is conscious from the third-person perspective, then there is no reason to suppose that we cannot imagine that something has a borderline case of consciousness in the same way. Of course, this is irrelevant to the modal status of borderline cases of consciousness, but the point is that there is no clear difference between imagining a clear case of consciousness and a borderline case of consciousness from the third-person perspective, so this cannot explain why we judge the first sort of case to be possible and the second impossible.

I think we should track the source of our intuition that consciousness that another person is, for instance, happy. If they are correct, then that would seem to contradict the statement in the text. However, what these theorists think is really going on is that there is a narrow base of more directly observable properties that our perceptual systems use to draw conclusions about someone's mental state. Thus, we can see that what is really at issue for the philosopher of perception is whether the inference about someone's mental state happens in perception or in cognition. Everyone agrees that conclusions about someone's mental state go beyond what is observable in a strict sense.

26 This is one way of thinking about Chalmers' zombie argument (Chalmers 1996). This fact is also instrumental in the arguments of Antony (2008) and Simon (2017) discussed above.
must be sharp back to our first-person acquaintance with consciousness. It is controversial what this acquaintance could amount to, but all I need are the following two claims. First, our experience of consciousness puts us in a position to know certain things about it. Second, one of the things we can come to know in this way is that consciousness must be sharp. Regarding the first claim, I should make it clear that I am not endorsing the controversial thesis of “revelation” (Johnston 1992) about conscious experience, on which merely bearing a phenomenal state is all that is needed to know its essence.\(^\text{27}\) I am instead endorsing the much weaker claim that we can sometimes know things about our conscious states just by undergoing them. I do not presently have an explanation for how this works, but denying that it occurs is a huge bullet to bite. A satisfying theory of consciousness should vindicate the claim that I can know an orange experience is more similar to a red experience than it is to a blue experience just by having the experiences in question. Only the most rabid of physicalists would deny this.\(^\text{28}\) My second claim is bound to be more controversial. The physicalist could grant that our experience of consciousness puts us in a position to know certain things about it while denying that its being sharp is one of those things. This is possible, of course, but I do not think the burden of argument at this point rests with

---

\(^{27}\)Consciousness not being vague is a modal fact because it concerns possible as well as actual cases. If modal facts and essence facts are connected in the way that Fine (1994) thinks they are, then perhaps I would have to claim that we can come to know something about a phenomenal experience’s essence just by having it. However, I would not have to claim that we can come to know a phenomenal experience’s complete essence in this way.\(^{28}\) Dennett (1988) might be an example.
me. The physicalist owes an explanation of why, if acquaintance can give us some insight into the nature of consciousness, we should draw the line at conclusions about its sharpness.

One might worry that I am illicitly privileging the first-person perspective here. Even if one thinks that we can come to know things about conscious experience just by having it, this does not rule out our coming to know things about it through third-person methods as well. If the situation is just that introspection suggests that consciousness cannot be vague, but our third-person investigations into consciousness suggest that consciousness must be vague, then perhaps we should suspend judgment rather than just going with what introspection tells us. I think this is misguided because, as I said before, we do not take anything that is available from the third-person perspective to be definitive of consciousness. Thus, third-person investigations into consciousness can only give us conditionals of the following form. If consciousness is correlated with observable property $P$, then $q$ conclusion about consciousness is true. Suppose that we have established some such conditional and we have established no other conditional that bears on $q$. If introspection tells us $\neg q$, then it would be wrong to say that the third-person evidence supports $q$. 21
2 Existing Physicalist Theories Imply that Consciousness Is Vague

In this section, I will consider two major categories of physicalist theories, functionalist theories\(^{29}\) and identity theories. I will argue that the theories in both categories, or at least the ones that have been offered so far, identify consciousness in vague terms. These categories do not cover all physicalist theories, but they do cover the most popular and influential physicalist positions. Furthermore, it should be clear that the argumentative strategy I use in this section can be generalized.

2.1 Top-Down vs. Bottom-Up Concepts

Before I begin, I wish to distinguish between what I will call bottom-up and top-down concepts. Bottom-up concepts are those that can be precisely defined in terms of fundamental physical, metaphysical, and logical/mathematical concepts. Because, as discussed above, fundamental properties must be perfectly sharp, bottom-up concepts will be perfectly sharp as well. Bottom-up concepts include the concepts employed by fundamental physics as well as those that are required to describe the fundamental metaphysical structure of the world. They also include high-level concepts built

\(^{29}\)Some functionalists might protest that their theories, although compatible with physicalism, are not physicalist \textit{per se} because they identify the state of being conscious not with any physical state but with a second-order state that could be realized by non-physical states if they existed. This is a fair point, but I will argue in this section that the second-order states themselves have borderline cases.
out of these ingredients by Boolean and quantificational means. Top-down concepts are those that can only be defined in non-fundamental terms. They include our rough-and-ready, everyday concepts, but they also include almost all of the concepts employed by the non-fundamental sciences. I claim that all top-down concepts are vague. For an example of a relatively precise top-down concept that nonetheless has borderline cases, consider the concept of being a water molecule. When hydrogen burns, hydrogen atoms combine with oxygen atoms to form water molecules. Imagine zooming in on a single instance of this process. The mereological fusion of the three atoms is not a water molecule at the beginning, and it is one at the end. Although very fast, this process is not instantaneous, so there will be moments in the middle when it is not clear whether the fusion has become a water molecule yet. Thus, the concept of being a water molecule has borderline cases. Even though chemistry is arguably the most precise of the non-fundamental sciences, the vast majority of chemical concepts do indeed have borderline cases. In the rest of this section, I will argue that the major contenders among physicalist theories of consciousness identify consciousness using top-down concepts.

30 Of course, both hydrogen and oxygen start as diatomic molecules themselves, but that detail will not be relevant here.
2.2 Functionalism

Functionalist theories identify being conscious with being in a state that fills a particular functional role. What functional role this is differs between theories, but the details will not be important for the purposes of this argument. What is important is the grain at which these theories define the role in question. Very broadly, functionalism seeks to define mental vocabulary in terms of how mental states respond to external stimuli, how they interact with one another, and what kind of behavior they produce. Consider how we might functionally describe pain. Pain is a state that is typically caused by bodily damage. Being in pain tends to cause many other mental states, including a desire to avoid the source of the pain. Depending on one’s other desires, being in pain might cause one to cry out or to withdraw from the painful stimulus. These clauses do not exhaust the functional role of pain, but they illustrate how a functionalist theory describes a network of states in abstract terms while remaining agnostic on the intrinsic natures of those states. A completed functionalist theory would be a full description of all the psychologically relevant aspects of such a network. The concepts that the functionalist then requires are those needed to describe external stimuli, like the concepts of having one’s body damaged or of seeing a car, those needed to describe behavior, like the concepts of crying out or of driving to the store, and those needed to describe the connections between mental states, stimuli, and behavior, particularly the concept of something’s typically causing
something else, and these concepts are all top-down.\textsuperscript{31}

I am not claiming that it is impossible for the functionalist to identify perfectly delineated, bottom-up replacements for these concepts. I discuss this move in section 3. Here, I am merely claiming that the concepts functionalists actually use are top-down, which is sufficient to show that the functionalist theories that have been offered so far all identify consciousness using vague concepts. If one thinks that causation is metaphysically fundamental, one might protest that the concept of something’s typically causing something else could be bottom-up. Perhaps the typically-causes relation is defined in terms of causation itself in some perfectly precise way. Of course, even if this were the case, it would not help the functionalist unless all the other top-down concepts were somehow eliminated from the theory, but I think it is instructive to look at why the typically-causes relation cannot be bottom-up. First, note that the typically-causes relation admits of exceptions that must be spelled out in top-down terms. My seeing a car typically causes me to believe that there is a car before me, and this fact is perfectly compatible with there being cases in which I see a car but am too distracted

\textsuperscript{31}Antony (2006b) offers four arguments for thinking that a functionalist theory of consciousness must be vague. His first two arguments are similar to the ones I make here. His third argument is addressed to the sort of functionalist who, like Lewis (1972), believes that a state just has to satisfy enough of our platitudes about consciousness in order to be consciousness. This sort of view renders consciousness vague because it is vague just how many platitudes is enough. His fourth argument relies on the incorrect assumption that if physical state $\Phi$ realizes functional role $F$, then a borderline case of $\Phi$ must be a borderline realizer of $F$. This is mistaken because a state could be a borderline case of $\Phi$ for reasons that do not affect its functional profile. However, his other three arguments stand.
to form the corresponding belief. This is because the concept of something typically causing something else contains implicit information about what sorts of exceptions are compatible with its still holding. The example just given should show that many of these exceptions would require top-down concepts to pick out explicitly. A further point to make is that, whether or not it is top-down, the concept of something’s typically causing something else is clearly vague. For example, does yawning typically cause other people to yawn? Sometimes it does, but there are plenty of people who are not particularly susceptible to contagious yawning. Also, if one is alone in one’s home, yawning will not typically have this effect. It is hard to know what to say in this case because it is a borderline case of the concept in question.

2.3 Identity Theories

Identity theories claim that the state of being conscious is identical to some kind of brain state, typically one picked out in neuro-physiological terms. These kinds of concepts might seem more precise than those involved in functionalist theories, but they too are top-down. Consider once more the example of pain. A common toy example of the identity theory assumes that being in pain is identical to having c-fibers that are firing. This is a concept that can be vague in several different ways. First, it can be clear that you have c-fiber-like things but vague as to whether they are c-fibers. Perhaps you have some neuro-degenerative condition that causes your c-fibers to decay. At the beginning of this process, it is clear that you have
c-fibers, and at the end, it is clear that you do not, but there will come a point somewhere in the middle when it is not clear whether what you have are badly decayed c-fibers or merely the remains of c-fibers. Second, it can be clear that there are c-fibers in your vicinity but vague as to whether you have them or not. Perhaps you are undergoing an experimental surgery to remove your c-fibers. Part way through the surgery, when your c-fibers are partially but not completely severed from the surrounding brain tissue, it will not be clear whether you really have those c-fibers. Third, it can be vague whether your c-fibers are firing. Perhaps you have some terrible sodium imbalance, so your c-fibers cannot quite fire properly. Depending on how badly they are malfunctioning, it might not be clear whether we should classify what they are doing as firing or not. What is at issue in these cases is not ignorance of the microphysics. Even if we knew the exact distribution of the fundamental physical properties, it would not settle the matter definitively. Thus, the concepts in question are top-down, and any physical state picked out by them must have borderline cases.\footnote{Antony (2006b) makes similar points.}

Again, I am not claiming that it is impossible for the identity theorist to find bottom-up replacements for neuro-physiological concepts. I am merely claiming that the identity theories that have been offered so far do not take this step, and thus they identify consciousness using concepts that have borderline cases.
3 Where Does This Leave Us?

3.1 Perfectly Precise Physicalism

It seems that the only option for the physicalist who does not want to rest content with merely gesturing toward consciousness is to look for a perfectly precise physical or functional property to identify with consciousness. Nothing I have said in this paper rules out the possibility of this project’s success. However, I think the physicalist who goes down this path will run into several issues.

A particular issue for an analytic functionalist is that a perfectly precise, bottom-up theory of consciousness will no longer be recognizably psychological. Analytic functionalism is the view that all we mean when we say that something is conscious is that it is in some state or other that plays the consciousness role. If the only way to spell out the functional role of consciousness is in fundamental physical terms, then it is not plausible that this is what we meant all along by the concept CONSCIOUS. Mastering the concept of consciousness does not require even implicit mastery of fundamental physical concepts.

A related issue that strikes at physicalism more broadly is that a bottom-up theory of consciousness would leave the physicalist without the ability to explain why consciousness is special.\textsuperscript{33} Suppose that the physicalist has

\textsuperscript{33}This is similar to Block’s “harder problem of consciousness,” although he casts the issue in epistemic terms (Block 2002).
identified a perfectly precise physical property that is plausibly identical to consciousness. The question then becomes, why should we care about this property in particular? The answer will not come from psychology or neuroscience because, as we have already established, they lack the resources to distinguish between the property in question and other very similar properties in its immediate vicinity. Even if this property is a possible precisification of some psychological or neuroscientific concept, there will be other precisifications that look equally good from the standpoint of those sciences. It will not have some special connection to our other mental states or to our behavior, because the property that is exactly similar to consciousness except with atom 357 moved one planck-length to the left will exhibit all the same connections to every discernible mental state and behavior. The answer will presumably not come from fundamental physics either because the property will just be one high-level compound property out of the great multitude that can be constructed using fundamental physical concepts. It seems that physicalists cannot call on a physical science that makes distinctions at the right grain in order to tell us why this property, consciousness, is so special. If they instead try to use first-personal features of this property to distinguish it from other physical properties in its vicinity, then they run the risk of lapsing into dualism.

This might seem, at first blush, to be just another statement of the explanatory gap between the physical and the phenomenal, but I think the issue I have identified here is different. The existence of the explanatory gap, as
described by Levine (1983), means that a physicalist cannot explain why the physical property that is supposed to be identical to a particular phenomenal property feels the way that it does. However, it does not prevent the physicalist from situating this physical property within our non-conscious cognitive system such that it exhibits all of the relational properties we would expect from that phenomenal property. This allows us to say that if any physical property is identical to the phenomenal property in question, it must be that one. To make this a little more definite, suppose that the physicalist claims that physical property Φ is identical to the property of seeing red. The explanatory gap means that she cannot explain why Φ has the particular phenomenal character that it does, but it leaves open the possibility that she can show that Φ is, for example, the unique physical property caused by the excitement of red cones and associated with memories of fire trucks. This would show at least that, among all other physical properties, Φ bears some special relation to seeing red. However, if what I am saying is correct, then the physicalist cannot even succeed at this more modest project because a different physical property, so similar to Φ as to be essentially indistinguishable, would have exactly the same relations to the excitement of red cones and memories of fire trucks. Similar remarks would apply to whatever physical property the physicalist wants to identify with being conscious.

The assumption that underlies this point is that, given a perfectly precise, high-level physical property Φ, we can find a different perfectly precise, high-
level physical property $\Phi'$ that plays an arbitrarily similar causal role.\textsuperscript{34} Of course, this might be false for fundamental physical properties because the fundamental laws could involve sharp cutoffs.\textsuperscript{35} If it is false for fundamental physical properties, then it is likely false for low-level derivative properties as well, like the property of being exactly 1g or 2g in mass. However, I think it is possible to maintain this assumption for high-level properties. The causal powers of derivative physical properties are constructed out of the causal powers of their fundamental constituents,\textsuperscript{36} and a high-level physical property presents us with enough knobs to turn that we can almost certainly find one which does not affect its causal role too much. Again, consider the case where $\Phi$ is a particular structural property. If $\Phi'$ is exactly the same except it involves one particle being located a vanishingly small distance to the left, then it seems clear that $\Phi$ and $\Phi'$ would have extremely similar causal powers.

One final issue for the physicalist is that, while most physicalists want to defer to sciences like psychology and neuroscience in the search for consciousness, this attitude of deference cannot be maintained while building a

\textsuperscript{34}If physics tells us that things are quantized at the fundamental level, then we may not be able to get arbitrarily close. If this is the case, then we can make due with the weaker assumption that $\Phi$ and $\Phi'$ play causal roles that are indistinguishable in psychological terms.

\textsuperscript{35}However, so far as we know, the fundamental laws are not like this, with the possible exception of quantum collapse.

\textsuperscript{36}I suppose a physicalist could believe that some physical laws directly pick out high-level properties, but I think most physicalists would reject this proposal, at least for consciousness, because it veers dangerously close to interactionist dualism. Most physicalists believe that no special laws are operative within conscious creatures.
bottom-up theory of consciousness. The arguments in this paper show that the resources available to non-fundamental sciences will never allow us to identify perfectly precise physical properties. Although the brain sciences might point us in the direction of consciousness, the physicalist will have to go beyond them in order to find a physical property that is even a candidate for being consciousness.

3.2 Panpsychism

Another sort of theory that my argument does not rule out is panpsychism. On this view, everything is conscious, from the fundamental particles to the whole of the universe. Panpsychism implies that there are no borderline cases of consciousness, then, because it does not attempt to draw a line anywhere. On most versions of panpsychism, the fundamental particles, or whatever the basic building blocks of reality turn out to be, bear the fundamental phenomenal properties, and the phenomenal states of higher-level objects, like human beings, are just combinations of the phenomenal states of their fundamental parts. There are physicalist versions of panpsychism, on which the fundamental physical properties are taken to be themselves phenomenal, so this could represent a way forward for the physicalist as well.\(^{37}\)

Two glaring issues have kept panpsychism outside the mainstream. First, it is deeply counterintuitive to think that every object is conscious. It is very hard to shake the belief that there is nothing it is like to be a rock or an

\(^{37}\)Philosophers who defend this sort of view include Stoljar (2001) and Strawson (2006).
electron. Second, it is not clear that separate centers of consciousness could combine to form a higher-level center of consciousness. A group of human minds, each with its own point of view, is not itself a single mind with a single point of view, or so we typically think. Why should the situation be any different with a group of minds that belong to electrons or quarks? Thus, although I grant that my argument leaves open panpsychist versions of physicalism, this will be cold comfort for most contemporary physicalists.

3.3 Dualism

The last sort of theory that my argument does not rule out is dualism. According to the dualist, some phenomenal properties are fundamental, and being conscious just amounts to bearing one of these fundamental phenomenal properties. Since it can never be vague whether a fundamental property is instantiated or not, it can never be vague whether a thing is conscious or not. However, most contemporary dualists believe that consciousness is still correlated with some physical property. The argument that I have given in this paper does show that whatever the dualist identifies as the physical correlate of consciousness cannot be vague.\(^\text{38}\) Thus, the dualist must engage in the same sort of project as the physicalist, namely, finding the perfectly precise physical property that is tokened when and only when consciousness is. Even so, I think the dualist is in a better position for two reasons. First, the

\(^{38}\) Antony (2006b) notes this as well. He claims that the dualist is no better off than the functionalist or the identity theorist, but I argue below that this is not the case.
dualist has no issue with saying why the physical correlate of consciousness, once she has discovered it, is special. It is special because it is correlated with consciousness, which is itself special due to its first-personal features. The physicalist rejects the claim that fundamental laws of nature pick out high-level physical properties that are only present in conscious creatures, which makes it difficult for her to explain why some of these properties are special. The dualist, however, welcomes this claim. The fundamental psychophysical laws do just that, and this is what allows the dualist to explain why the high-level physical properties in question are special. Second, the dualist is less likely to balk at the idea of going beyond the domain of psychology and neuroscience in identifying the physical correlate of consciousness.

Conclusion

It can never be vague whether something is conscious. To deny this would be deeply counterintuitive. However, most physicalist theories identify consciousness using concepts characterized in non-fundamental terms. High-level concepts like these will always have possible borderline cases, so these theories cannot tell us what consciousness really is. At best, the concepts of psychology and neuroscience can point us in the direction of consciousness, but they cannot have the last word on its physical basis. The physical basis of consciousness must be sharp-edged, and thus we will need tools from fundamental physics or metaphysics to identify it.
References


35


