IS CONSCIOUSNESS VAGUE?

GEOFFREY HALL

ABSTRACT. This paper defends the thesis that possibly something is such that it is borderline whether it is phenomenally conscious. Consciousness is vague. There are a number of arguments that purport to show the opposite view that consciousness is precise. I argue that these arguments depend on principles that are inconsistent with plausible, and somewhat standard, principles governing the notion of a borderline case.

Is consciousness vague? This paper will argue that it is. But first, we need to get clear on the meaning of the question.

I will take vagueness to consist in the possibility of borderline cases. *Being F* is *vague* if and only if possibly something is such that it is borderline whether it is F; otherwise it is *precise*.[1] Thus *being red* is vague because it is possible that something is such that it is borderline whether it is red and *being a prime number* is precise because it is not possible that something is such that it is borderline whether it is prime. More to the point, consciousness is vague if and only if it is possible that something is such that it is borderline whether it is conscious. To figure out whether consciousness is vague we must figure out whether it has borderline cases.[2]

The sense of 'consciousness' at issue is *phenomenal consciousness*. One is phenomenally conscious if and only if there is something it is like—something it *feels* like—to be one.[3] So if consciousness is vague, in the intended sense of 'consciousness', possibly it is borderline whether some creature has an associated phenomenology: it is borderline whether there is something that it is like to be that

---

[1]Any discussion of vagueness must take some stand on how talk of borderline cases is to be regimented. This paper takes as primitive the sentential operator *it is borderline whether*, though I do not mean to take any stand on whether this operator is perfectly natural or anything of that sort. It is perhaps more standard to take as primitive the predicate 'is borderline', reflecting the widely held view that vagueness is ultimately a linguistic matter. But the choice between these two approaches is mostly verbal: there are well known strategies, successful up to a point, for translating between discourse that uses sentential operators and discourse that uses a predicate of sentences, for a classic discussion see Kaplan 1986. More on topic, Dorr (2009) argues the the choice between the two approaches is ultimately one of priority. I prefer the use of a sentential operator mostly because it facilitates less messy formulations of principles that involve quantifying in. For some further reasons to taken sentential operators as primitive see Bacon 2018, ch. 4.

[2]Thus according to this view, bordelineness is more fundamental than vagueness and precision. While this is a pretty standard assumption, some deny it. For a defense of the opposite view, that vagueness and precision are more fundamental than borderliness, see Bacon 2018, ch. 12.

[3]See Nagel 1974.

creature.[4] Suppose, to take one example, that it is borderline whether ants are conscious. Then it will be borderline whether there is something it is like to be an ant. The question, "What is it like to be an ant?" will lack any determinate, or definite, answer.

Many philosophers have thought that consciousness just couldn't be vague. On their view, consciousness is precise: conscious beings are determinately conscious and non conscious beings are determinately not conscious.[5] Those sympathetic to this view often analogize consciousness to a light. While a light can be more or less bright, it is either on or it is off. Even the dimmest light is on. So too, they say, with consciousness. There are cases in which someone is barely conscious. But even a barely conscious agent is a conscious agent.[6]

But the thesis that consciousness is precise has some significant theoretical costs. Antony (2006) argues that if consciousness is precise, functionalism, identity theory and even certain forms of dualism must be false. O'Rourke (2017) argues that if consciousness is precise, physicalism is false. Perkins and Bayne (2013) argue that if consciousness is precise, representationalism is false. These arguments taken together make a strong theoretical case for the vagueness of consciousness. If these arguments stand up, the vagueness of consciousness would be a consequence of many of our best theories of consciousness.

What has convinced philosophers to think otherwise? While there is a lot of variation in the arguments that have been given, many of them have a similar flavor: if consciousness were vague, we should be able to find some clear cases of something such that it is borderline whether it is conscious. But we can't. So it isn't. Most authors who have pushed this line have not taken note of the verificationist assumptions underlying it. On the proposed view, borderline cases occupy a portion of reality that is fully within our view. With enough investigation, we could decide all questions about what was borderline and why. I see no reason to believe this. In fact there are positive reasons not to since it can be shown, given standard principles governing borderlineness, that there are some borderline cases that cannot be known.

---

[4]Note that it could be that all borderline cases of consciousness are non-agents. Assuming the law of excluded middle, everything is either conscious or it is not. But it could be that everything that is borderline conscious is not conscious.

[5]We can define 'determinately' in terms of borderlineness by stipulating that it is determinate that φ if and only if φ and it is not borderline whether φ.

[6]While remarks like these are common among those who hold that consciousness is precise, on my view, they fail to express that view altogether. It is a consequence of classical logic that everything is either conscious or it is not conscious just as it is a consequence of classical logic that everything is either red or not red. But it doesn't follow from this that *being red* is precise, since there are cases in which something is a borderline case of 'red'. In other words, on any classical theory of vagueness, like supervaluationism or epistemicism, for instance, the fact that something is either conscious or not has no bearing on whether consciousness is vague.

Now it may be that there is some genuine insight behind these arguments that is obscured in their formulation. The basic question is this: is there some data that the assumption that consciousness is precise can better account for then the assumption that consciousness is vague? I will answer that there isn't: all of the data that people have appealed to in order to support that precision of consciousness can be accommodated within a theory that permits borderline cases of consciousness. Since this theory is—or at least is not obviously not—consistent with our best theories of consciousness, this provides strong reasons to prefer it over its alternative.

This paper comes in four sections. In §1, I address a family of arguments for the precision of consciousness that have been given in the literature and argue that they depend on premises that conflict with plausible and widely accepted principles governing borderlineness. In §2, I show how some recent work by Bacon (2018) can be used to provide a version of these arguments that avoids the objections I provide in §1. But this argument, I argue, is too strong, as it can be used to similarly argue for the precision of notions that are clearly vague. In §3, I take a step back and ask more generally whether the supposed precision of consciousness accounts for some data that cannot be accounted for on the assumption that consciousness is vague. I argue that it does not. Finally in §4 I argue that in the absence of a convincing argument for the precision of consciousness, we ought to accept that it is vague.

While I won't assume any specific theory of borderlineness in what follows, I am going to assume classic logic. Since the majority of this paper will consist of defending the view that consciousness is vague against objections, it seems to me an appropriate assumption to make. Many of our leading theories of vagueness today are classical theories. And for good reason: classical logic, construed as a body of truths, is an extremely strong and simple theory. If the arguments against the vagueness of consciousness require us to reject that theory, that in itself would be a cost.

1. The Argument from Recognition

If there are borderline cases of consciousness, where are they? When we attempt to identify a clear case of something that is such that it is borderline whether it is conscious, we seem to come up short. Of course, it may be that, say, ants are borderline conscious. But we do not know this because for all we know, it is determinate that they are conscious. If there are any cases in which a thing is a borderline case of consciousness, we don't know of them. Indeed, it is somewhat difficult to *imagine* what a borderline case of consciousness would be like.

Can these sorts of considerations be assembled into a convincing argument for the precision of consciousness? Two recent attempts to do so can be found in Antony 2006 and Simon 2017. Both articles emphasize the connection between competence with vague concepts and some sort of sensitivity to the borderline cases of that concept. As Antony says:

> [C]ompetence with a vague concept **F** entails a sensitivity to the location and character of **F**'s borderline regions, and such sensitivity is typically expressed by manifesting categorizing dispositions in and around **F**'s borderline regions such as: increased hesitation about whether to judge individuals as F or not-F, denying both that an individual is F and that it is not-F, changing one's mind as to whether it is F or not-F, judging it to be neither clearly F nor clearly not-F.(2006, p. 242)

He goes on to further endorse a connection between the ability to have thoughts involving a vague concept, and the ability to conceive of borderline cases of a vague concept:

> [G]iven that **F** is vague, competence with the concept... require being able to conceive ... borderline Fs.(2006, pp. 250-251)

Antony then argues that we are not able to conceive of cases in which it is borderline whether something is conscious, and so our concept of consciousness must be precise.[7]

Simon (2017) also thinks that competence with some vague concept requires sensitivity to its borderline cases. His initial explanation of this appeals to a putative principle concerning our ability to *recognize* borderline cases:

> [W]hen something is a borderline case of a predicate, there is a *positive characterization* of that borderline case–a way of filling in the details of the case that will show any sufficiently competent speaker what makes it a borderline case, and accordingly put any such speaker in a position to see why it is a borderline case.(2017, p. 2106)

He goes on to provide a more detailed characterization of the thesis, explaining key terms like "positive characterization" along the way. This thesis is then used to argue that consciousness is precise since, given the conceivability of zombies—creatures physically and functionally like us but lacking consciousness—we are never in a position to provide a positive characterization of a borderline case of consciousness.

---

[7]At this point in the paper I am being deliberately a bit loose as to whether it is words, concepts or properties that are the primary bearers of vagueness. The reason is that the claims I am interested in can all be formulated with the borderlineness operator, use of which is neutral as to the priority question.

I think that with Simon's first pass formulation of the positive characterization claim, we can already see that something is amiss. Suppose that it is borderline whether φ. Then if Simon's thesis is correct, competent speakers must be able to be in a position to see why it is borderline whether φ, and thus be able to be put in a position to know that it is borderline whether φ. Call that thesis *verificationism*.

VERIFICATIONISM: If it is borderline whether φ, then someone is able to be put in
a position to know that it is borderline whether φ.

One immediate worry with VERIFICATIONISM is that it looks to entail that vagueness depends on the existence of agents. However it's not obviously false that vagueness preceded the existence of agents. If so, then there were borderline cases without any agents, let alone agents in a position to know about such borderline cases. I don't want to push this line of objection too far though. I suspect that with a little bit of tweaking, we could avoid it altogether while still capturing the main idea that Simon is getting at in the above quoted passage. There are deeper worries for VERIFICATIONISM.

A widely accepted principle in theorizing about vagueness is that borderline cases preclude knowledge.[8] While it is a consequence of both epistemicism and supervaluationism, it can also be motivated by appealing to premises that appear to be mundane truths together with out background assumption that logic is classical. Suppose a given individual is a borderline case of 'rich'. Intuitively, in such cases, one does not know this individual is rich nor does one know that the individual is not rich. But since, by the law of excluded middle, the individual is either rich or not rich, classical logic guarantees that either the individual is rich and we do not know this, or the individual is not rich and we do not know this. Now this kind of ignorance in borderline cases is a hallmark of the epistemicist view. [9] But it also a crucial part of the supervaluationist view. If knowledge entails truth, and truth is supertruth, then knowledge precludes borderlineness, since borderlineness precludes supertruth.[10] More generally:

PRECLUSION: If one is in a position to know that φ, then it is not borderline whether
φ.

---

[8]While widely accepted, there are some people who deny this. In particular both Barnett (2011) and Dorr (2003) argue that it is false.

[9]See Williamson (1994).

[10]See Keefe 2000, ch. 8 for more on the idea of truth as supertruth.

The problem is now this: VERIFICATIONISM and PRECLUSION jointly entail that whenever it is borderline whether φ, it is not borderline whether it is borderline whether φ. That is, they entail that there is always a determinate fact as to whether something is a borderline case:

DETERMINACY OF BORDERLINENESS: If it is borderline whether φ, then it is determinate that it is borderline whether φ.

For suppose that it is borderline whether φ. By VERIFICATIONISM, someone is in a position to know that it is borderline whether φ. And so by PRECLUSION it is not borderline whether it is borderline whether φ. Thus by the definition of determinacy, since it is borderline whether φ but not borderline whether it is borderline, it follows that it is determinate that it is borderline whether φ.

The DETERMINACY OF BORDERLINENESS is false. I was once a child, but now I am not. Thus on pain of contradiction, there is some first moment at which I stopped being a child. [11] What moment was this? It seems impossible to say. The standard explanation for why this is is that the moment occurred during a period of time in which it is borderline whether I was a child. Cases like these are in fact how the notion of borderlineness is often introduced. The explanation assumes that it was once the case that it was borderline whether I was a child. Since I am no longer a borderline case of a *child*, on pain of contradiction, there must be some first moment at which it stopped being borderline whether I was a child. What moment was this? Again it seems impossible to say. And it seems impossible to say for the same reasons that it was impossible to say when I stopped being a child *simpliciter*. This suggests that we should give the same explanation: the reason we cannot say which moment it stopped being borderline whether I was a child is that it occurred during a period of time in which it was borderline whether it was borderline whether I was a child. In particular, there were some moments before it stopped being borderline whether I was a child in which it was borderline whether I was a child, and it was borderline whether it was borderline whether I was a child. This provides us with a counterinstance of the DETERMINACY OF BORDERLINENESS.

The phenomenon of higher-order vagueness also calls into question Antony's principle connecting vague concepts with our ability to conceive of borderline cases, albeit in a slightly more subtle way.

---

[11]Note that here is where the assumption of classical logic comes in. Classic logic entails that vague predicates have *cutoffs*: if one goes from being a child to a nonchild, then since at every moment one is either a child or nonchild, there is some first moment at which one is a nonchild and was a child a moment before.

To illustrate, suppose that it is both borderline whether Thomas is bald and borderline whether it is borderline whether Thomas is bald. Then while we might not be able to know, in this case, that it is borderline whether Thomas is bald, there will nevertheless be some other possible situation in which Thomas is determinately borderline bald. There is thus no bar to our conceiving of that possible situation. What is needed is a vague sentence that not only admits of borderline cases, but further does *not* admit of determinate borderline cases. Let φ be a sentence satisfying the following conditions:

(1) Possibly it is borderline whether φ

(2) Necessarily if it is borderline whether φ, it is borderline whether it is borderline whether φ.

If a sentence can satisfy both of these conditions, it would be vague, since it is possibly borderline. But it isn't clear to me that we would be able to *conceive* of the cases in which it was borderline. Or at least, if we could conceive of them, it is unclear whether we could be in a position to *know* that we were conceiving of them. Suppose for instance that I asked you to conceive of a person who was bald, but was such that it was borderline whether they were bald. If you were to attempt to conceive of this by explicitly imagining a person and what they looked like, you would be in no position to *know* whether the individual you are imagining was both bald and borderline bald: the imagined individual being borderline bald would preclude anyone from knowing, in the imagined situation, that the individual was bald. Thus it would also appear to preclude *you* from knowing that the individual in question was a borderline case of 'baldness'. More generally, if there were a sentence φ that was possibly borderline but never determinately so, whatever instance of its borderlineness one attempted to imagine would be an instance that, for all one knows, isn't an instance of borderliness.

Such cases also cast doubt on Antony's more general contention that competence with a borderline concept requires sensitivity to its borderline regions. If it is always borderline where those borderline regions are, it may it may be that the borderline regions in question are one's that competent speakers simply cannot be sensitive to in the way that Antony supposes. For suppose it is borderline whether φ and borderline whether it is borderline whether φ: if someone asks me whether φ and I begin to hesitate in a way indicative of vagueness, it would be natural to express my hesitation by saying "it's borderline." But supposing that knowledge is the norm of assertion, so that one should assert only what one knows, this assertion would not be apt in the context at

hand, given PRECLUSION: I'm not in a position to know it's borderline, and so am not in a position to assert that it is.

I think this shows that *if* there are sentences satisfying (1) and (2), then we ought to reject the idea that competence with a vague predicate requires sensitivity to its borderline regions. Do any sentences satisfy the joint constraints of (1) and (2)? We could demonstrate the *consistency* of that assumption by showing that there are models, belonging to a class of models whose logic is plausibly the joint logic of possibility and borderlineness, on which both (1) and (2) come out true. But there are also plausible examples.

Suppose that φ is a sentence such that ⌜it is borderline whether φ⌝ and ⌜it is borderline whether it is borderline whether φ⌝ are both true (not above I argued such cases can exist on the grounds that our ignorance of when I stopped being a borderline case of *child*). Then one possible candidate sentence satisfying (1) and (2) is the sentence ⌜*actually* φ⌝. That is, plausibly we will both have that ⌜possibly it is borderline whether actually φ⌝ is true and that ⌜necessarily if it is borderline whether actually φ, then it is borderline whether it is borderline whether actually φ⌝ is true. So for example, consider some x such that it is borderline whether x is a child, and borderline whether it is borderline whether x is a child. Then plausibly it will be necessary that it is borderline whether it is borderline whether x is actually a child, where 'actually' is being used to express the modal operator that always shifts the world of evaluation back to the actual world.

There are some principles we could posit governing actuality that entail that the sentence ⌜actually φ⌝ has the desired features. For instance:

> It is borderline whether φ if and only if necessarily it is borderline whether actually φ.

I find this principle pretty intuitive given the intended interpretation of the actuality operator. Note that it is trivial that it is borderline whether φ if and only if necessarily actually it is borderline whether φ, given the intended interpretation of 'actually'. We could thus derive the above principle from the principle that that borderlineness and actuality necessarily commute:

> Necessarily actually it is borderline whether φ if and only if it is borderline whether actually φ.

This strikes me as at least initially plausible, though not uncontroversial. One might reject it on grounds like the following. Suppose that φ is determinately true at the actual world w but false at some other world v. Suppose moreover that it is borderline whether $w = v$. That is, some

8

distinct world is such that it is borderline whether it is the actual world. Perhaps then we should say that while it is determinate that φ, it is borderline whether actually φ. On this view it is not determinately the case that φ if and only if actually φ.

This picture appeals to borderline identities. Many authors have thought that borderline identity is impossible, often citing Evans (1978). But as Bacon (2018) has recently pointed out, Evan's doesn't quite establish that conclusion. Since it is never borderline whether $a = a$, the assumption that it is borderline whether $a = b$ entails, given Leibniz's law, that it is not the case that $a = b$. But for all that argument shows, it is indeed borderline whether $a = b$.

For my purposes, we needn't decide whether or not such borderline identities exist. For suppose that one rejects my possible example of sentence satisfying (1) and (2) on the grounds that appeal to borderline identities. This will then provide us a different example of a sentence satisfying (1) and (2). For suppose that it is borderline whether $a = b$. Then it is possibly borderline. So the sentence '$a = b$' satisfies (1). Evan's proof shows, moreover, that it is determinate that if $a = b$, it is determinate that $a = b$. But one can prove that insofar as 'it is determinate that' is a factive normal modal operator, the statement 'it is determinate that: $a = b$ only if it is determinate that $a = b$' implies the statement 'it is borderline whether $a = b$ only if it is borderline whether it is borderline whether $a = b$'. [12] Thus if there are vague identities, they provide instances of *both* (1) and (2).[13] Either way, we have some grounds to be skeptical that there is any strong connection between competence with a vague predicate and the ability to conceive of borderline cases of that predicate. Some borderline cases are completely hidden from view, and so competence with the relevant predicate needn't involve any kind of sensitivity to those cases.

---

[12]*Proof sketch.* Let $\mathcal{L}_\Delta$ be a propositional language equipped with a normal modal operator Δ expressing determinacy. Let $\nabla\varphi$ abbreviate $(\varphi \wedge \neg\Delta\varphi) \vee (\neg\varphi \wedge \neg\Delta\neg\varphi)$. Intuitively $\nabla\varphi$ expresses that it is borderline whether φ. Let Λ be a normal modal logic extending the normal modal logic KT. We want to show that $\Delta(\varphi \to \Delta\varphi) \vdash_\Lambda (\nabla\varphi \to \nabla\nabla\varphi)$ (i.e., that $(\Delta(\varphi \to \Delta\varphi) \to (\nabla\varphi \to \nabla\nabla\varphi)$ is a theorem of Λ). It suffices to show that this is a theorem of KT and so suffices, in turn, to show that it is valid on every reflexive frame. So let $\mathcal{F}$ be an arbitrary reflexive frame, $\mathcal{M}$ a model based on $\mathcal{F}$ and $w \in W$ a point in $\mathcal{M}$. Suppose that $\mathcal{M}, w \Vdash \Delta(\varphi \to \Delta\varphi)$ and $\mathcal{M}, w \Vdash \nabla\varphi$. We want to show that $\mathcal{M}, w \Vdash \nabla\nabla\varphi$. The formula $\Delta\varphi \to \neg\nabla\varphi$ is valid on every reflexive frame. And so Since $\mathcal{M}, w \Vdash \varphi \to \Delta\varphi$ (by the reflexivity of $\mathcal{F}$), $\mathcal{M}, w \nVdash \varphi$. Hence $\mathcal{M}, w \Vdash (\neg\varphi \wedge \neg\Delta\neg\varphi)$ (by the supposition that $\mathcal{M}, w \Vdash \nabla\varphi$). So there is some v accesible from w such that $\mathcal{M}, v \Vdash \varphi$. Since $\mathcal{M}, w \Vdash \Delta(\varphi \to \Delta\varphi)$, $\mathcal{M}, v \Vdash \varphi \to \Delta\varphi$ and so $\mathcal{M}, v \Vdash \Delta\varphi$. Thus $\mathcal{M}, v \nVdash \nabla\varphi$ (appealing again to the fact that $\Delta\varphi \to \neg\nabla\varphi$ is valid on reflexive frames). Thus $\mathcal{M}, w \nVdash \Delta\nabla\varphi$. Hence $\mathcal{M}, w \Vdash \neg\Delta\nabla\varphi$. Therefore $\mathcal{M}, w \Vdash (\nabla\varphi \wedge \neg\Delta\nabla\varphi)$ and so $\mathcal{M}, w \Vdash \nabla\nabla\varphi$. Since $\mathcal{M}$ and w were chosen arbitrarily this suffices to show the desired result.

[13]As Lewis points out, this proof is only convincing in logics that include Leibniz's law, in its schematic form. Supervaluationists have some grounds for rejecting that principle. I'm not a supervaluationist though so this hardly shows the response doesn't work.

The precision of consciousness cannot be *derived* from our failure to find borderline cases. To think otherwise involves making unjustified verificationist assumptions about agents' relation to borderline cases. In the section after next, I will ask whether our putative inability to find borderline cases nevertheless makes it probable that consciousness is precise. But before doing that, I am going to address a further argument for the precision of consciousness that while in some ways analogous to Simon and Antony's arguments, sidesteps some of the problematic verificationist assumptions they make.

2. The Argument from Supervenience

Both Antony (2006) and Simon (2017) articulate principles that give voice to the thought that vagueness is insubstantial in some sense. For Antony, this means that that competence with a vague predicate requires the ability to conceive of borderline cases. The vagueness of a predicate couldn't be something we fail to notice in theorizing with it. Competence with vague predicates requires some sort of recognition of that vagueness. The fact that the predicate is vague cannot be a "deep fact" that we are completely insensitive to in our theorizing. Similarly, Simon argues that borderline cases can always be positively described in a way that reveals them to be borderline cases. Again, the borderline cases are required to be something competent speakers are sensitive to. That x is a borderline case of 'F' cannot be a deep fact, beyond the grasp of ordinary speakers.

This sort of denial of deep facts about vagueness is crucial to the success of their arguments that consciousness is not vague. But as we have seen, these principles look to conflict with certain widely accepted theses concerning higher-order vagueness and the relationship between knowledge and vagueness. For this reason, I think their arguments fail. However I think there is another possible argument from the supposed insubstantial nature of vagueness to the precision of consciousness that avoids the pitfalls of their arguments. The central premise of the argument is a principle that connects our *credences* in the precise with our credence in the vague that aims to capture the idea that vagueness is in some sense a "shallow" phenomenon. In this section I will formulate this argument and explain how it avoids some of the problems that the arguments of Antony (2006) and Simon (2017) faced. Ultimately, though, I think this argument fails as well. I will show that a parallel argument can be constructed from Bacon's principle to the conclusion that certain uncontroversially vague notions are precise. This shows that something is amiss with the initial argument, without coming to any definite conclusions on just what that something is.

2.1. **The Argument from Rational Supervenience.** There may be several different ways to formulate a principle connecting our credences in vague phenomenon with our credences in precise phenomenon. For definiteness, I will present a version of this principle that has recently been defended by Bacon (2018). The most straightforward formulation of Bacon's principle assumes the controversial view that it is propositions, as opposed to sentences, that are the fundamental bearers of vagueness and precision. For simplicity I'll assume that view for the remainder of this section, leaving open whether and how it can be reformulated in the framework that takes sentences to be the fundamental bearers of vagueness and precision.

Let's start with the principle:

RATIONAL SUPERVENIENCE: If Cr and Cr' are two conceptually coherent ur-priors, then if $Cr(p) = Cr'(p)$ for any precise proposition p, then $Cr(p) = Cr'(p)$ for *any* proposition p.[14]

If RATIONAL SUPERVENIENCE is true, then any two non-conceptually confused agents with the same evidence whose credences agree on how things are *precisely*, that is with precise matters, will agree on how things are *simplicter*. In particular, they will agree on all vague matters. To take one example, suppose that two rational agents' credences are distributed exactly the same way over the number and distribution of hairs on someones head (for simplicity we'll assume that such facts are precise, though in reality they of course exhibit vagueness as well). Then insofar as they have the same evidence, and they are not making any conceptual mistakes, they will then be equally confident in whether the individual in question is bald.

Notice that the principle RATIONAL SUPERVENIENCE differs from the principles proposed by Antony (2006) and Simon (2017) by focusing on the distinction between vagueness and precision, rather than on the distinction between borderlineness and determinacy. Even if it is determinate that A is bald, it is still vague whether A is bald because it is possibly borderline whether A is bald. Vagueness as I am understanding it is merely the possibility of borderlineness, and so is compatible with determinacy. For our purposes, we can say that a proposition is precise if and only if it is not vague.[15]

An ur-prior Cr is "conceptually coherent" just in case an agent whose credences are represented by Cr is not thereby conceptually confused. For the purposes of the argument I will take the notion

---

[14](Bacon , 2018, p. 161).

[15]To be clear, this modal account of vagueness and precision is one Bacon (2018) rejects. However, nothing much hangs on this issue here so for continuity with the previous section I will continue to implicitly define vagueness and precision in terms of borderliness.

of conceptual confusion as primitive. However I think it is plausibly tied to conceivability in the following way: if a proposition p is conceivable, then there is some conceptually coherent ur-prior Cr such that $Cr(p) = 1$. Conceivable propositions are propositions for which an agent could be certain of without thereby being conceptually confused in any way. This principle represents an important bridge between the principle of RATIONAL SUPERVENIENCE and the thesis that zombies are conceivable. The argument for the precision of consciousness will make use of both theses.

The argument proceeds by *reductio*. So suppose, looking for a contradiction, that consciousness is vague and let C be the proposition that something is conscious. Let P be the conjunction of all physical truths. Plausibly, if consciousness is vague, then all phenomenal features are vague as well. The specific phenomenal features consist of consciously representing, being in pain and the like. Many people already believe that these features are vague, and so the assumption that they are conditional on consciousness being vague seems to me minimal. Then, supposing that zombies are conceivable, we have:

(1) It is conceivable that $P \wedge \neg C$.

So give the proposed connection between conceivability and conceptually coherent ur-priors, we have

(2) There is a conceptually coherent ur-prior Cr_1 such that $Cr_1(P \wedge \neg C) = 1$

Of course, since $P \wedge C$ is actually true, there should also be a conceptually coherent ur-prior such that Cr_2 such that $Cr_2(P \wedge C) = 1$. That is

(3) There is some rationally coherent ur-prior Cr_2 such that $Cr_2(P \wedge C) = 1$.

Now plausibly, all *non-phenomenal* truths supervene on the physical truths, in the sense that if two rationally coherent ur-priors agree on the physical truths, they ought to agree on the chemical truths, biological truths and so on. (I take it that two rationally coherent ur-priors will trivially agree on the necessary truths and so will also agree on the mathematical truths, though strictly all that we are required to assume is that we can choose Cr_1 and Cr_2 so that they agree on the necessary truths, which seems reasonable). So since Cr_1 and Cr_2 both agree on the physical truths, we have

(4) For any non-phenomenal proposition p, $Cr_1(p) = Cr_2(p)$

But by assumption all phenomenal truths are vague, and so the precise propositions are completely included in the non-phenomenal propositions. Thus from (4) it follows that

(5) For any precise proposition p, $Cr_1(p) = Cr_2(p)$

And so from (5) and RATIONAL SUPERVENIENCE we have:

(6) For any proposition p, $Cr_1(p) = Cr_2(p)$

But (6) contradicts (2) and (3), since Cr_1 and Cr_2 are conceptually coherent ur-priors Cr_1 we must have $Cr_1(C) = 0$ and $Cr_2(C) = 1$.

In brief, the argument proceeds by noting a tension between the following three theses: (i) agreement on precise propositions induces agreement on vague propositions, (ii) agreement on the physical propositions *does not* induce agreement on phenomenal propositions, and (iii) there are no precise phenomenal propositions. The basic thought is this: if we agree on the vague if we agree on the precise, and all physical propositions are precise, then it should follow that agreement on the physical propositions induces agreement on the phenomenal ones. But it doesn't, as zombie cases show.

2.2. **Rational Supervenience, Indeterminate Perception and Spectrum Inversion.** How compelling is the argument from RATIONAL SUPERVENIENCE? Not very, it turns out, for we can construct a parallel argument from RATIONAL SUPERVENIENCE to the conclusion that certain uncontroversially vague properties are precise. Thus we can be certain that at least one premise of the the argument from RATIONAL SUPERVENIENCE is false without having to say which one it is. Whereas the argument from RATIONAL SUPERVENIENCE depended on the assumption that zombies were conceivable, the parallel argument depends on the conceivability of certain sorts of spectrum inversion cases in which the inverted subjects are experiencing *indeterminate perception*. Before describing the case I will briefly review these phenomena.

Let's start with specrum inversion. Let C be the set of phenomenal color properties, i.e. properties that characterize what it is like to see red, see blue and so on. We can imagine such properties as arranged in a hue circle, corresponding to the hue circle characterizing the actual colors. Let $\sigma : C \to C$ be a permutation of these properties that "inverts" them, mapping properties like *sensing redly* to *sensing greenly* and *sensing bluely* to *sensing yellowly*.[16] Many have thought that it is conceptually coherent to suppose that there are beings x and y that are functionally and physically alike, except that for any property F in C, x has F if and only if y has σF. That is, many have thought it is coherent to suppose that there be spectrum inversion without physical

---

[16]For more details see Byrne (2020).

or functional difference. For our purposes, it suffices to note that *if* one accepts the conceivability of zombies, one should probably also accept the conceivability of spectrum inversion. Two agents spectrum inverted with respect to one another are much more similar than any conscious being is to a nonconsious being. This philosophical zombies represent a much more radical case than spectrum inversion.

If two agents are spectrum inverted with respect to one another, then an object will appear red to the one if and only if it appears green to the other. Now in ordinary cases in which an object appears red, it also appears F, for some determinate F of *red*. However as Hardin (1985, 1988) has emphasized in his well known critique of sense data theory, there are cases of *indeterminate color perception*: cases in which an object appears to have some color without appearing to having any determinants of that color. In particular, when an object occurs in peripheral vision, our visual system will attribute determinable properties to objects without attributing their more specific determinates. Thus an object might appear red, without appearing crimson, scarlet, pink etc. This raises the possibility of two spectrum inverted agents who differ *merely* in that to one, a certain object appears red, whereas to the other it appears green. Consider the following case.

An individual, Nonvert, experiences central vision loss, but retains peripheral vision. In the periphery of Nonvert's vision is exactly one object o that appears to be red. Since o is in the periphery of Nonvert's vision, o appears to be red without appearing to be F for any determinate F of *red*. Nothing else appears to be any way to Nonvert. Across town there is another individual, Invert, that is a physical and functional duplicate of Nonvert, but is spectrum inverted with respect to Nonvert. Like Nonvert, Invert experiences central vision loss. Moreover, invert is living in a qualitative duplicate of Nonverts environment. There is thus an object o' qualitatively indistinguishable from o that appears within Inverts peripheral vision. Since Invert is spectrum inverted with respect to Nonvert, o' appears to Invert to be green. But since o' is in the periphery of Inverts vision, it does not appear to Invert that o' is F, for any determinate F of *green*.

In the described situation, Invert and Nonvert differ in their phenomenal properties. Nonvert senses redly whereas Invert senses greenly. But these are the *only* phenomenal properties they differ over. Since they are physical and functional duplicates, we can suppose that they are emotionally alike, things all appear to have the same shapes etc. They are both conscious, since they both sense Fly, for some F. Thus Nonvert and Invert differ in their phenomenal properties, but they

do not differ in their *precise* phenomenal properties, since sensing redly and sensing greenly are uncontroversially vague. It seems to me that *if* we are prepared to accept zombie scenarios as being conceivable, we should also regard this scenario as conceivable. In the proposed scenario, there isn't anything particularly questionable about Nonvert: central vision loss is not merely conceivable but actual. Moreover, the sort of spectrum inversion in the proposed scenario seems much less radical than the idea of functional and physical duplicates lacking in phenomenal properties altogether.

So suppose we grant that this is a conceivable scenario. We can then argue as follows. Let P be the conjunction of all of the physical propositions true in the above scenario. Let R be the proposition that there is some x such that it appears to Nonvert that x is red. Finally, let C be the conjunction of all phenomenal properties that Nonvert has apart from the property of being an x such that it appears to x that o is red. Then the conjunction $P \wedge C \wedge R$ is conceivable, since this just describes how Nonvert actually is in the above scenario. So there is some conceptually coherent ur-prior Cr_1 such that $Cr(P \wedge C \wedge R) = 1$. But it is also conceivable that things stand with Nonvert phenomenally just as they stand with Invert: it is conceivable that Nonvert has just the phenomenal properties attributed by C, and just the physical properties properties attributed by P, and yet lack the phenomenal property attributed by R. Thus $P \wedge C \wedge \neg R$ is also conceivable. So there is a conceptually coherent ur-prior Cr_2 such that $Cr_2(P \wedge C \wedge \neg R) = 1$. Since Cr_1 and Cr_2 differ over some proposition, they must then differ over some precise proposition, by RATIONAL SUPERVENIENCE. But they agree on all of the physical propositions, and agree on all of the phenomenal propositions other than R. The only proposition over which they differ is R. It follows then, from RATIONAL SUPERVENIENCE that R is precise. But since R is uncontroversially vague, this argument leads to a conclusion that is uncontroversially false.

One might respond that while R has a vague *content*, it doesn't follow that R itself is vague. They might point to the above argument as *grounds* for holding that R is in fact precise. [17] But as Morrison (2012) has argued, it is a plausible hypothesis that the similarity structure of color experiences reflects the similarity structure of the colors themselves. So since colors themselves exhibit vagueness, it seems plausible to suppose that color experiences would exhibit vagueness in the same way, provided that they mirror the colors in their similarity structure.

---

[17]Thanks to an anonymous reviewer for raising this point.

I think this argument suffices to show that the argument from Rational Supervenience fails. It also illustrates a tension in the whole approach to arguing for the precision of consciousness on the basis of the supposed insubstantial nature of vagueness. The basic approach says that vagueness is never a deep fact, but if consciousness were vague, then there would be deep facts about vagueness. There are different ways to spell that idea out. But no matter how it is spelled out, is hard to see how one can avoid the parallel argument: given the conceivability of spectrum inversion in cases of indeterminate perception, there must be vague propositions that are in some sense "inscrutable" from the precise facts if we grant that whether a thing is conscious is inscrutable from the precise facts. The property of being conscious is not special in that respect. I think this shows that ultimately there is no way that this line of reasoning can succeed.

3. Why are there no clear borderline cases of consciousness?

With natural languages, vagueness is the default state. Almost every word we use is vague. Exceptions tend to be logical and mathematical vocabulary, and perhaps certain vocabulary from fundamental physics. Even in these domains, it is often a major theoretical achievement when one singles out some precise meaning for the word in question, as when precise definitions of continuity and computability were found. It is difficult to avoid speaking vaguely, and often takes theoretical work to speak precisely.

This suggest that our initial assumption ought to be that 'phenomenal consciousness' is vague. Oftentimes the only associated description is that a being is phenomenal conscious if and only if there is something it is like to be that being. Such a description doesn't obviously precisely pick out any quality in the way that, say, the truth table for 'and' precisely defines the conjunction operation.

I think there are two main reasons people have resisted this sort of "anti-exceptionalist" view of phenomenal consciousness. One is that it might be thought to have uncomfortable moral implications. For if there were beings who were borderline conscious, there would be beings whose moral status may be borderline. But many have thought one cannot have borderline moral status. While I think this is an interesting argument, for reasons of space, I will have to set it aside for this paper. I want to look at the second main reason for resisting the anti-exceptionalist thesis that conscsiousness is vague. We can formulate this reason as a simple argument:

(P1) There are no clear cases in which it is borderline whether something is conscious.

(P2) The best explanation of (P1) is that consciousness is precise.

(C) Therefore, consciousness is precise.

If consciousness is precise, then nothing can be borderline conscious. Necessarily, if something is conscious, it's determinate that it is conscious. And necessarily, if something is not conscious, it's determinate that it is not conscious. If both of these principles are true, we of course have a straightforward explanation as to why no clear borderline cases of consciousness can be found. No borderline cases exist and so *ipso facto* no clear borderline cases exist.

Suppose we grant premise (P1). There aren't any clear borderline cases of consciousness. Still, we shouldn't infer that consciousness is precise. For there are other explanations available for why no clear borderline cases can be found that make more minimal assumption. I'll consider two possible explanations. The explanations are not necessarily incompatible with one another. But they approach the topic from slightly different viewpoints.

The first possible explanation is one that embraces the existence of borderline cases of consciousness, but denies the existence of determinate borderline cases of consciousness. This theory thus embraces both of the following principles:

(i) Being conscious is vague.

(ii) Necessarily, for any x, if it is borderline whether x is conscious, it is borderline whether it is borderline whether x is conscious.

Principle (ii) suffices to explain why there are no clear borderline cases of consciousness. If x was clearly borderline conscious, then we should be in a position to know that x is a borderline case. But if we are in a position to know, then x is not a borderline borderline case of consciousness. Thus from (ii), we can infer that there are no *clear* borderline cases of consciousness.

Note that the thesis that consciousness is precise logically entails (ii). If consciousness is precise then necessarily for any x, it is not borderline whether x is conscious. But (ii) is compatible with (i). Thus (ii) is strictly weaker than the thesis that consciousness is precise. What this shows is that we can account for the truth of (P1) with logically weaker assumptions than the argument supposes. This by itself undercuts the force of the inference to the best explanation somewhat.

This response might appear a bit *ad hoc*, however. What reason, other than to explain (P1), is there to suppose that no determinate borderline cases of consciousness exist? Recall though that many have thought that our best theories of consciousness entail that it is vague. So if our

starting point is one of these theories, what this response shows is that it would be and overreaction to *abandon* such a theory because we cannot find clear borderline cases of consciousness. There are other possible explanations for the hiddenness of borderline cases of consciousness that don't require us to abandon our antecedent theory of consciousness.

Even so I think there is a less *ad hoc* explanation as to why we cannot clearly identify borderline cases of consciousness. Suppose that it is borderline whether x is conscious. Now if x is not conscious, then plausibly x cannot know that it is borderline whether x is conscious: such knowledge would have to be occurrent rather than dispositional. But unconscious things lack occurent mental states altogether. Now suppose that x is conscious. Since x is borderline conscious, they are not *determinately* in some occurent mental state. And so if present knowledge that one is conscious is always determinate, it would be at best borderline whether x knows that it's borderline whether x is conscious. Thus any borderline case of consciousness would be an agent for which it is at best borderline whether they know they are a borderline case.

Can anyone other than x know that it's borderline borderline whether x is conscious? If zombie's are conceivable, it is hard to see how they could know for certain. Whatever evidence they have that x is borderline conscious will be *a priori* compatible with x's being determinately conscious. But if we grant that the publicly available evidence fails to conclusively establish borderline cases, then we are able to explain why borderline cases of consciousness are hard to find. We cannot find them from within, at least not determinately. And we cannot locate them from the outside—at least not with certainty. Thus if there are borderline cases of consciousness, we should *expect* them to be elusive. The inference from the hiddenness of borderline case of consciousness to their nonexistence is thus called into question.

4. Conclusion

We've seen that several recent arguments for the precision of consciousness fail. Since the fact that consciousness is vague is predicted by our best theories of consciousness, I think we should tentatively conclude that consciousness is vague. This conclusion is very tentative: there are several other arguments for the precision of consciousness that I have not addressed. I think ultimately such arguments are not persuasive, but I have not argued for that here. Less tentatively, arguments that appeal to the idea that vagueness must be "scrutable" from the precise facts are bound to fail.

If they were successful, analogous arguments would show that uncontroversially vague notions are precise. The proponent of the view that consciousness is precise is in need of a new approach.

REFERENCES

Antony, Michael 2006. Vagueness and the Metaphysics of Consciousness, *Philosophical Studies* 128: 515-538.

Antony, Michael 2008. Are our Concepts CONSCIOUS STATE and CONSCIOUS CREATURE Vague?, *Synthese* 68/2: 239-263.

Bacon, Andrew 2018. *Vagueness and Thought*. Oxford University Press.

Barnett, David 2011. Does Vagueness Exclude Knowledge?, *Philosophy and Phenomenological Research*, 82/1: 22-45.

Dorr, Cian 2003. Vagueness with Ignorance, *Philosophical Perspectives*, 17/1: 83-113.

Dorr, Cian 2009. Iterating Definiteness, in *Cuts and Clouds: Vagueness, its Nature and its Logic*, ed. Sebastiano Moruzzi and Richard Dietz, Oxford: Oxford University Press: 550-575.

Evans, Gareth 1978. Can There Be Vague Objects? *Analysis* 38: 208.

Keefe, Rosanna 2000. *Theories of Vagueness*. Cambridge University Press.

Kaplan, David. 1986 Opacity, in *The Philosophy of W.V. Quine*, eds. Lewis Edwin and Paul Arthur Schilpp Open Court: 229-289.

MacFarlane, John 2020. Book Review: *Vagueness and Thought*, *Philosophical Review* 129/1: 153-158.

Morrison, John 2012. Color in a Physical World: A Problem Due to Visual Noise, *Mind* 121: 333-73.

Nagel, Thomas 1974. What is it like to be a bat? *Philosophical Review*, 83: 435-456.

O'Rourke, Joshua. 2017. *Property Dualism as a Solution to the Mind-Body Problem*. Ph.D. thesis, Princeton: Princeton University. URL = http://arks.princeton.edu/ark:/88435/dsp019019s5231

Perkins, Ryan and Tim Bayne 2013. Representationalism and the Problem of Vagueness, *Philosophical Studies* 162/1: 71-86.

Simon, Jonathan 2017. Vagueness and Zombies: Why 'Phenomenally Conscious' Has no Borderline Cases, *Philosophical Studies* 174: 2105-2123.

Williamson, Timothy. 1994. *Vagueness*. Routledge.