The Gradability of 'Conscious'

* Forthcoming in *Philosophy & Phenomenological Research*

ANDREW Y. LEE | University of Toronto, Philosophy
POPPY MANKOWITZ | University of Bristol, Philosophy

* Authors are listed in alphabetical order by last name. All authors contributed equally.

Abstract

Are some creatures "more conscious" than others? A number of consciousness researchers have aimed to answer this question. Yet some have claimed that this question doesn't even make sense. They claim that 'conscious' (in the phenomenal sense) never occurs as a *gradable adjective,* meaning an adjective that permits degree expressions ('more F than', 'slightly F', etc.) and that's associated with a degreed property. Both sides face an explanatory burden: they must explain why some competent speakers seem confused about the meaning of 'conscious'. We argue that the question does make sense: 'conscious' sometimes functions as a minimal-standard gradable adjective. But we will also explain why some theorists have been skeptical about gradable uses of 'conscious'. Along the way, we address the objection that many gradable constructions involving 'what it's like' expressions are infelicitous, distinguish two interpretations of 'phenomenal consciousness', and discuss how our semantic arguments bear on the metaphysical question of whether consciousness comes in degrees.

WORD COUNT: 11,047 (main text + footnotes)

KEYWORDS: phenomenal consciousness, gradability, gradable adjectives, degrees of consciousness, more conscious, subjectivity, phenomenal character, what it's like

Introduction

Consider the following question:

> MORE CONSCIOUS?
> Are some creatures more conscious than others?

On the face of it, this is a metaphysical question asking whether the property of consciousness has a degreed structure.[1] But it also raises a semantic question, of whether the term 'conscious' is a *gradable adjective.* Gradable adjectives—such as 'tall', 'open', and 'intelligent'—are adjectives that permit degree expressions ('more F than', 'slightly F', etc.) and that are associated with degreed properties (such as height, openness, and intelligence). A negative answer to the semantic question means that the metaphysical question doesn't even make sense: the phrase 'more conscious than' would not even be meaningful. The main goal of this paper is to evaluate the semantic question, though we will also say a few things about the metaphysical question.[2]

There's no doubt that some senses of 'conscious' are gradable. But the question of this paper is whether the *phenomenal* sense of 'conscious'—where an entity counts as 'conscious' just in case there is something it's like to be that entity—is one of them. If 'conscious' (hereafter in the phenomenal sense) is gradable, then sentences where 'conscious' is used with a degree expression (which we'll hereafter call 'target sentences') are felicitous. As examples of such target sentences, consider the following:

[1] An analogous question arises for state (as opposed to creature) consciousness. We will mostly frame the discussion in terms of creature consciousness, though our arguments will likewise apply to state consciousness.

[2] See Lee [2023] for more extensive discussion of the metaphysical question.

Target Sentences

(1*a*) Humans are more conscious than fish.

(1*b*) If a snail is conscious, then it's probably only a little conscious.

(1*c*) AIs may soon be slightly conscious.

(1*d*) Perhaps psychedelic states are more conscious than sober states.

(1*e*) If we stimulate your prefrontal cortex, then we hypothesize that you'll become more conscious than you were before.

A number of philosophers have contended that these sorts of sentences are infelicitous or incoherent.[3] After all—they might argue—it seems questionable to say, 'There's more of what it's like to be *x* than there is of what it's like to be *y*'. Perhaps the gradability of non-phenomenal senses of 'conscious' yields an illusion that the phenomenal sense of 'conscious' is gradable too. Or perhaps these sentences involve some sort of coercion, where the gradable environment forces a non-standard interpretation of 'conscious'.

However, the above sorts of sentences are commonly asserted by consciousness researchers, including in contexts where it's clear that phenomenal consciousness is the relevant subject matter. Many investigations of consciousness proceed on the assumption that some creatures are more conscious than others, and some theories of consciousness explicitly build degrees into the structure of consciousness. If it turns out that 'conscious' isn't gradable, then one might wonder how to interpret the utterances of these speakers. It would be surprising if a significant proportion of experts are systematically misusing one of their central theoretical terms.

The aim of this paper is to adjudicate this dispute by defending the following answer to the semantic question:

[3] As examples, see Bayne, Hohwy, & Owen [2016: 408], Birch, Schnell, & Clayton [2020: 790], Carruthers [2020: 23], and Tye [forthcoming: 8].

GRADABILITY

The phenomenal sense of 'conscious' *is* a gradable adjective when used by (at least some) researchers.

Along the way, we will develop several novel arguments concerning the semantics of 'conscious'.

Our arguments will leave open whether 'conscious' is *always* used as a gradable adjective. Instead, we aim merely to argue that 'conscious' (in the phenomenal sense) is *sometimes* used as a gradable adjective, especially by some consciousness researchers. In other words, we will argue that some occurrences of 'conscious' have a literal meaning that involves a scale measuring degrees of consciousness. In this respect, our thesis is modest. Nevertheless, it's dialectically significant. As noted above, some authors have explicitly rejected GRADABILITY, and have appealed to this rejection to dismiss any theory of consciousness that postulates degrees of consciousness. By arguing that 'conscious' sometimes functions as a gradable adjective, this paper resists this strategy of dismissal.

The question of the gradability of 'conscious' parallels recently discussed questions about the gradability of other philosophical terms, such as 'possible', 'justified', 'rational', 'true', and 'know'.[4] Although the focus of this paper is on 'conscious', many of our arguments could be generalized to other debates where the gradability of a term is under dispute.

Here's the structure of the paper: **§1** identifies the relevant sense of 'conscious' and discusses our methodology; **§2** argues that it's unlikely that the target sentences involve non-phenomenal senses of 'conscious'; **§3** argues against other strategies for explaining away gradable uses of 'conscious'; **§4** counters an objection to GRADABILITY that appeals to 'what it's like' expressions; **§5** argues that 'conscious', when used as a gradable adjective, functions as a *minimal* gradable adjective; **§6** disentangles two ways of

[4] See Lassiter [2017] on 'possible', Hawthorne & Logins [2021] on 'justified', Siscoe [2022] on 'rational', Mankowitz [2023] on 'true', and Pavese [2017] on 'know'.

interpreting 'phenomenal consciousness', one of which permits gradable uses more naturally than the other; §7 discusses how our semantic arguments bear on the metaphysical question of whether consciousness comes in degrees.

§1 Gradability

The standard methods for evaluating whether an adjective is gradable are to see whether it naturally occurs in *comparative constructions* ('more F than', 'as F as', etc.) and with *degree modifiers* ('slightly F', 'completely F', etc.).

Comparative Constructions

(2*a*) That walrus is bigger than the other one.
(2*b*) The dirt road is bumpier than the paved road.
(2*c*) Aldo's towel is drier than Bezawit's shirt.
(2*d*) This soup bowl is less full now than it was a minute ago.
(2*e*) ? This number is more (even / prime / finite) than the other one.

Degree Modifiers

(2*f*) That's a very big walrus.
(2*g*) The dirt road is slightly bumpy.
(2*h*) Aldo's towel is completely dry.
(2*i*) This soup bowl is half full.
(2*j*) ? That's a (very / slightly / completely / half) (even / prime / finite) number.

To test whether degree expressions occur naturally with a given adjective, it's common to consult linguistic intuitions of competent speakers and to

look at patterns of use in *corpora*, meaning collections of real-world spoken or written expressions. But it's tricky to apply these tests to 'conscious'.[5]

Most uses of 'conscious' in the phenomenal sense occur in theoretical contexts. In ordinary contexts, it's often difficult to identify whether 'conscious' is being used in the phenomenal sense or some other sense.[6] This makes it unobvious whether the linguistic intuitions of ordinary speakers track the phenomenal sense of 'conscious' (versus some other sense). Furthermore, since there's no straightforward way to isolate uses of 'conscious' in the phenomenal sense in ordinary contexts, conducting searches of large corpora would fail to generate reliable data about the phenomenal sense of 'conscious'.[7] Because of this, we will mostly set aside considerations about how ordinary speakers use 'conscious' in everyday contexts, and we will focus on uses by consciousness researchers, who are able to isolate the phenomenal sense of 'conscious'. This may raise the concern of whether it makes sense to investigate the semantics of a theoretical term, given that the meanings of theoretical terms are often stipulated. However, even theoretical terms exhibit systematic semantic behaviors. Furthermore, the meanings of many theoretical terms—especially ones that have a rich

[5] To our knowledge, the only prior attempt at applying these tests to 'conscious' occurs in Brogaard [ms]. However, many of the examples involve occurrences of 'conscious' with a prepositional complement, which is usually taken to express the awareness sense of 'conscious' (e.g., 'He was considerably less conscious *of the stimuli* when it was extremely low intensity'). Since it's controversial how the awareness sense of 'conscious' relates to the phenomenal sense, it's difficult to assess whether these examples genuinely support the claim that 'conscious' in the phenomenal sense is a gradable adjective. See §2 for discussion of various senses of 'conscious'.

[6] There's some uncertainty as to how often the phenomenal sense of 'conscious' is clearly expressed in ordinary contexts. Perhaps when ordinary speakers use 'conscious', they typically don't have one particular sense in mind.

[7] For example, a search of the iWeb Corpus (Davies 2018) for 'more conscious than' reveals results like 'Young guys these days are even *more conscious* than women regarding their looks' and 'He shows himself even *more conscious* than they of the grandeur and holiness of God'. Such examples, which clearly involve non-phenomenal senses of 'conscious', would need to be manually excluded by individually inspecting each entry.

history of use and that play important roles in our theories—aren't determined purely through stipulations.

Many consciousness researchers find the target sentences felicitous. Indeed, we can construct a small 'corpus' of examples where real theorists seem to be using 'conscious' in the phenomenal sense with degree expressions (which we have italicized):

(3*a*) "[A]re newborn babies conscious, and to what extent? Are […] some animals *more conscious* than others?" (Tononi 2004, *p*.2)

(3*b*) ("[W]e can use this […] to make a hypothesis about phenomenology. Babies are *more conscious* than we are." (Gopnik 2007, *p*.504)

(3*c*) "Any system whose functional connectivity and architecture yield a Φ value greater than zero has at least a trifle of experience […] So by that measure, a fly would be *less conscious* than you are in your deep sleep." (Koch 2011, *p*.132.)

(3*d*) "It's arguable that systems with a *minimal degree of consciousness* (ants?) have only a minimal degree of moral status." (Chalmers 2022, *p*.340)

(3*e*) "It may be that today's large neural networks are *slightly conscious*." (Sutskever 2022)

(3*f*) "I would be quite willing to ascribe *very small amounts of degree* [of consciousness] to a wide range of systems, including animals." (Bostrom 2023)

(3*g*) "[E]ven if it turns out that humans are *more conscious* than octopuses on some dimensions while octopuses are *more conscious* than humans on other dimensions, it may still be the case that both are *more conscious* (simpliciter) than fish." (Lee 2023, *p*.7)

These examples indicate that some theorists use the phenomenal sense of 'conscious' as if it were a gradable adjective. One conclusion would be that this is because the phenomenal sense of 'conscious' *is* a gradable adjective, at least when used by these theorists. This would be to endorse GRADABILITY. On the other hand, some theorists have reported to us that they find these sentences infelicitous, even when focusing on the phenomenal sense of 'conscious'. So, another conclusion would be that we should interpret these examples with care and avoid taking them at face value.

This means that both proponents and opponents of GRADABILITY face an explanatory burden. For the proponent of GRADABILITY, the challenge is to explain why the target sentences sound infelicitous to some speakers. For the opponent of GRADABILITY, the challenge is to explain why the target sentences sound felicitous to, and are used by, some theorists.

Both sides could attempt to meet the challenge by questioning the reliability of their opponent's verdicts about felicity. Perhaps the target sentences are moving beyond ordinary uses of 'conscious' in a way that prevents linguistic intuitions from providing a clear verdict.[8] Or perhaps the verdicts of theorists on the other side are shaped by which theory of consciousness they favor.[9] But since these considerations can cut both ways, the arguments in this paper won't essentially hinge on the reliability of either side's verdicts. While a few of our arguments will appeal to tests targeting linguistic intuitions, most will rely on broader semantic and philosophical arguments.

The rest of the paper makes a case for GRADABILITY. We will start by exploring the kinds of explanations that opponents of GRADABILITY could appeal to and argue that none of them work. In addition, we will respond

[8] A standard view is that linguistic intuitions are the surfacing of tacit knowledge of one's language. These intuitions are shaped by ordinary uses of language. As one moves beyond everyday occurrences of expressions (see *fn*.6), one's tacit knowledge of English may not provide clear verdicts on which expressions are felicitous.

[9] See Lee [2023] on what various theories of consciousness entail about degrees of consciousness.

to a few further objections that might be raised against GRADABILITY. Then we will suggest some explanations for why the target sentences strike some speakers as infelicitous.

§2 Polysemy

The word 'conscious' is *polysemous*: it's associated with several distinct yet related senses.[10] An opponent of GRADABILITY might therefore argue that all felicitous 'more conscious' sentences involve a non-phenomenal sense of 'conscious'. In what follows, we will identify other senses of 'conscious' and explain why it's unlikely that the target sentences involve these other senses.

Intentional Action

'Conscious' is sometimes used to mean that an action (whether mental or behavioral) was performed deliberately, as opposed to automatically.[11] This sense is normally expressed using an adverb ('I *consciously* decided to take the scenic route'), or an adjective that is syntactically combined with an action-denoting expression ('My *decision* to take the scenic route was *conscious*'). By contrast, in the target sentences, it's expressions with denotations consisting of individuals or states—not actions—that combine with 'conscious'.

[10] We remain neutral about the correct analysis of polysemy. One traditional approach that's compatible with our discussion holds that polysemous expressions have multiple *characters,* where these are functions from contexts of utterance to *content* (the information expressed by a particular occurrence of an expression; see Kaplan [1989]).

[11] OXFORD ENGLISH DICTIONARY: 'consciousness' (intentional action sense) $=_{def}$ done or created deliberately; aware of what one is doing or intending to do; having a purpose and intention in one's actions ('There is need for a conscious decision'; '[T]hrough no conscious effort I was…happier than I had been since childhood').

Awareness

'Conscious' is sometimes used to denote a relation between a subject and an object.[12] In these uses, 'conscious' has a complement consisting of a prepositional phrase ('He was conscious *of the time*') or a clause ('He was conscious *that time was running out*'). When used in this way, 'conscious' is roughly synonymous with 'aware' ('He was aware of the time' / 'He was aware that time was running out').[13] By contrast, in the target sentences listed earlier, the term 'conscious' lacks an overt complement. Since the awareness sense of 'conscious' requires an object of awareness (supplied by the adjective's complement), the target sentences don't appear to express this sense of 'conscious'. Furthermore, while 'conscious' in the awareness sense isn't typically used to denote a relation between a *state* and an object, some of the target sentences concern states ('Perhaps psychedelic states are more conscious than sober states').

It's worth noting that there are cases where an adjective that normally occurs with a complement is interpretable even when it occurs without an overt one. For example, 'aware', 'afraid' and 'allergic' normally occur with a complement specifying the object of awareness, fear or allergy ('He's (aware of / afraid of / allergic to) the dog'), but may sometimes be used without a complement ('He's (aware / afraid / allergic)'). But in such cases, the context (or sometimes prior discourse) supplies a suitable object.[14] If no suitable object is supplied, then the sentence sounds incomplete; it's natural to react with the question 'He's aware *of what*?'.

---

[12] OXFORD ENGLISH DICTIONARY: 'consciousness' (awareness sense) $=_{def}$ having knowledge or awareness; able to perceive or experience something; often used with a clause or prepositional phrase (e.g., 'I was suddenly conscious that I was jabbing my finger at him'; 'Vaguely conscious of the transiency and instability of material life').

[13] See Silva & Siscoe [2024] for a recent analysis of 'aware [of the fact] that' expressions.

[14] We focus here on expressions that require a definite or indexical implicit argument, as opposed to an existential one (see Lasersohn [1993] and Condoravdi & Gawron [1996]). We suspect that 'aware' and the awareness sense of 'conscious' belong to this category.

It's implausible to hold that interpreters consistently assume a context rich enough to supply a suitable object for the target sentences. For example, if the sentence 'AIs may soon be slightly conscious' is interpreted in the same way as 'AIs may soon be slightly aware', it's not obvious what the object of awareness is supposed to be. And it doesn't seem natural for someone to react with the question 'AIs may soon be slightly conscious *of what*?'. This makes it unlikely that the target sentences are expressing the awareness sense of 'conscious'.

Wakefulness

'Conscious' is sometimes used to mean wakefulness or responsiveness to the environment ('She slipped out of consciousness and began dreaming').[15] This sense of 'conscious' uncontroversially allows for gradable constructions: one individual can be more awake or responsive than another. However, some target sentences—for example, 'Perhaps even a dreaming human is more conscious than an awake fish'—cannot be interpreted as expressing the wakefulness sense.

Intelligence

Some uses of the target sentences seem intended to convey the idea that the relevant entities are more intelligent (or more cognitively sophisticated in some amorphous sense). We are skeptical that this is a genuine sense of 'conscious': no dictionary we examined lists it as such, and we suspect these sorts of uses involve conflation between different psychological attributes. Furthermore, some of the target sentences ascribe consciousness to states (as opposed to creatures), but it's hard to know what it would mean for, say, psychedelic experiences to be more intelligent than sober experiences.

[15] OXFORD ENGLISH DICTIONARY: 'consciousness' (responsiveness sense) $=_{def}$ aware of and responding to one's surroundings; having one's mental faculties in an active and waking state. ('He was seen within a short time of wounding and was fully conscious without evidence of intracranial damage').

Finally, it's also possible to explicitly exclude this kind of interpretation: for example, consider the sentence 'Future AIs will be highly intelligent, but not even slightly conscious'.

Other Senses

There are other senses of 'conscious', such as being especially aware of one's own appearance or attributes, having mastery over one's mind, and being attentive to social issues (being "woke"). These senses of 'conscious' are obviously irrelevant to the theoretical debates within which the target sentences occur: scientists are not debating whether fish are "woke."

This illustrates a more general point: it's unlikely that our target speakers—namely, experts who research phenomenal consciousness—are conflating 'conscious' in the phenomenal sense with 'conscious' in one of the other senses. For example, when Gopnik [2007] suggests that babies are 'more conscious than we are', she's clear that she intends to make a claim about phenomenal consciousness. It's implausible that she's mistakenly expressing the claim that babies are more intelligent, more awake, more aware of some contextually salient object, or more intentional in their actions than adults.

§3 Non-Standard Uses

An opponent of GRADABILITY owes an explanation of why the target sentences sound felicitous to and are used by some theorists. We have just argued that this explanatory burden cannot be met merely by appealing to polysemy. Another natural approach is to appeal to *non-standard uses* that enable the sentence to convey something different from its standard semantic meaning.[16] In cases involving non-standard uses, interpreters are able to make sense of sentences where non-gradable adjectives are used with degree expressions.

[16] We remain neutral on whether the non-standard content is pragmatically conveyed or is semantically expressed via some non-standard literal meaning.

Such an approach has some initial plausibility. Adjectives such as 'dead', 'illegal', 'married', and 'nuclear' are normally classified as non-gradable: they occur with degree expressions neither frequently nor naturally, and it's hard to make sense of the target properties as coming in degrees. Yet each is nevertheless sometimes used with degree expressions:[17]

(4*a*) My patients are more dead than your patients.
(4*b*) What Aldo did is very illegal.
(4*c*) France's energy supply is partly nuclear.
(4*d*) Aldo is more married than Bezawit.

To explain these occurrences of non-gradable adjectives with degree expressions, three kinds of non-standard use are commonly distinguished in the literature:[18]

Coercion
The use provides a scale that doesn't contribute to its standard meaning but is still related to the standard meaning.

Quantification
The adjective's nominal arguments provide a domain of individuals or parts that can be evaluated with respect to the property denoted by the adjective's ordinary meaning.

Metalinguistic Comparison
The context provides a scale that measures aptness of expressions or content.

[17] See Burnett [2017: 45, 113] for more discussion of these sorts of examples.
[18] See Mankowitz [2023: 97].

Each kind of non-standard use makes a scale or domain available (but not a scale supplied by the adjective's standard meaning), and the degree expression is then interpreted relative to that scale or domain. In what follows, we will explain each kind of non-standard use in more detail—focusing on sentences 4*a*–*d*—and argue that none of them yield plausible diagnoses of the intended meanings of the target sentences.

§3.1 Coercion

Consider the sentence 'My patients are more dead than your patients'. This might be used to convey the meaning expressed by 'My patients are further (temporally) from the point of having died than your patients are'. Even though 'dead' is non-gradable, its use in the comparative construction generates a new scale associated with distance from having reached the point of death. As another example, the sentence 'What Aldo did is very illegal' might be used to convey the meaning expressed by 'What Aldo did is an illegal act associated with a very severe punishment'.

These uses involve *coercion*, or cases where an attempt at combining an expression with another expression that has the wrong type of meaning leads to an interpretation different from the standard one.[19] For present purposes, our concern is with cases where using a degree expression with a non-gradable adjective causes the adjective to be associated with a scale. Coercion requires a context where interpreters can recover a scale related to the ordinary meaning of the non-gradable adjective. Often, the most natural scales measure how close things are to attaining the non-degreed property, or how far things are from having attained it.

It's implausible, however, to interpret the target sentences as coerced. For example, when someone says 'Humans are more conscious than fish', they don't mean that humans are closer, in some sense, to attaining the non-degreed property of being conscious than fish are, or that humans

[19] Coercion is a general phenomenon that applies to non-adjectival phrases as well. See Pylkkänen & McElree [2006].

are further, in some sense, from the point of having attained that property than fish are. In fact, it's not even clear what scale would be naturally recovered if consciousness doesn't come in degrees. The most natural candidates might be scales that measure how closely something resembles entities that are conscious. But this doesn't yield natural coerced interpretations of our target sentences.

§3.2 Quantification

The sentence 'My patients are more dead than your patients' might also be used to convey the meaning expressed by 'More of my patients are dead than your patients'. Similarly, the sentence 'France's energy supply is partly nuclear' might be used to convey the meaning expressed by 'Some of France's energy supply is nuclear.'

These uses are *quantified*, in that the adjective's nominal arguments provide a domain and the degree expression is interpreted as a quantifier on that domain. Such uses require a degree expression that can be matched with a related determiner[20]—for instance, 'completely', 'a little', 'half', 'mostly' and 'more'—and a nominal argument with a meaning that provides a bounded domain of individuals or parts. For example, 'my patients' and 'your patients' might provide a domain of patients, and the degree expression 'more' might be interpreted as the quantifier denoted by the related determiner 'more'. By contrast, a quantified use would *not* be possible for 'The bacterium is very dead': 'the bacterium' does not provide a natural domain, and 'very' cannot be matched with a related determiner ('very of the individuals …'?).

[20] As examples, there are related determiners for 'completely / totally' ('all (of the)'), 'a little' ('some / a few (of the)') and 'more' ('more (of the)'); but there are no related determiners for 'perfectly', 'very', etc. It's difficult to clearly state the nature of this link, but it seems to involve semantic meanings that invoke similar "amounts." There may also be lexical constraints; for example, 'completely' and 'perfectly' are semantically equivalent degree modifiers, but only the former allows a quantified interpretation.

It's implausible, however, to interpret the target sentences as quantified. For example, when people say 'If a snail is conscious, then it's probably only a little conscious', they clearly don't mean that only a small proportion of parts of the snail are conscious. When people say 'Humans are more conscious than fish', they don't mean that more humans (or parts of humans) are conscious than fish (or parts of fish). Furthermore, some target sentences don't even include the types of degree expressions that are required to license quantified uses. For instance, quantified uses are not possible for 'AIs may soon be slightly conscious', because the degree expression 'slightly' cannot be matched with a related quantifier.

§3.3 Metalinguistic Comparison

Finally, the sentence 'My patients are more dead than your patients' might be used to convey the meaning expressed by 'It's more apt to say that my patients are dead than it is to say that your patients are dead'. As another example, 'Aldo is more married than Bezawit' might be used to convey that it is more apt to say that Aldo is married than to say that Bezawit is married. Metalinguistic uses are possible only with an occurrence of 'more', 'less', or 'as much', and when the context supplies a scale measuring aptness.[21] For instance, 'The patients are completely dead' cannot be understood as metalinguistic, because it doesn't involve the right sort of comparative construction.

Many uses of the target sentences cannot be analyzed as metalinguistic comparatives. First, many of them lack the appropriate degree expressions ('more', 'less', or 'as much'). Second, even when there is an appropriate comparative construction, the metalinguistic interpretation is often implausible. For example, when people say 'Perhaps psychedelic states are more conscious than sober states', they don't mean that it's more apt to

[21] See Morzycki [2011: 40-41, 70]. Scales that measure aptness have been analyzed by means of degrees of imprecision (according to Morzycki) or of preference (according to Giannakidou & Yoon 2011).

say that psychedelic states are conscious than to say that sober states are conscious. Instead, it seems equally apt to say that either are conscious, since both kinds of states are determinately conscious and known to be conscious.

§3.4 General Remarks

We will end by mentioning some general reasons for doubting that all occurrences of the phenomenal sense of 'conscious' could be explained in terms of non-standard uses. First, as noted previously, there are some target sentences where the constraints on non-standard uses preclude some such interpretations. For example, it's not possible to interpret 'AIs may soon be slightly conscious' as quantified or as metalinguistic, due to the degree expression and nominal involved. Second, the opponent of GRADABILITY incurs the burden of providing a plausible reinterpretation of every use of a target sentence.[22] That's a significant burden: even if some target sentences may be explained by appeal to non-standard uses, it seems unlikely that all will be amenable to this kind of analysis. Third, our target uses occur in contexts where researchers are aiming to express the intended ideas in a precise manner. But non-standard uses are more common in informal, everyday speech, since they rely on interpreters to reconstruct what a speaker might have intended to convey.

We conclude that appealing to non-standard uses isn't a promising strategy for the opponent of GRADABILITY to explain why some theorists find the target sentences felicitous. And we argued in §2 that appeals to polysemy are also unpromising. We will next turn to some objections to

[22] For a similar observation, see Lassiter [2017: 91, 132]. As Rob Stainton has pointed out to us (p.c.), the opponent of GRADABILITY doesn't need to provide the same type of reinterpretation (e.g., a quantified one) for all uses of all target sentences. Even so, the burden remains significant, because there are plausibly some occurrences of sentences that are not amenable to any of the reinterpretation strategies, especially in cases where a researcher is aiming to express their idea in a precise manner.

GRADABILITY, and we will explain why they aren't good grounds for rejecting the thesis.

§4 'What It's Like'

Some have resisted GRADABILITY by appealing to cognate expressions, such as 'what it's like', 'subjective point of view' and 'being conscious'. Since there's no straightforward way of replacing 'conscious' in the target sentences with these expressions while preserving acceptability, some theorists have taken this to indicate that there's something defective about the target sentences. Furthermore, it may seem that the denotations of those expressions—what it's likeness, subjective points of view, and being conscious—cannot come in degrees.

Although we haven't seen this objection developed in detail, several authors have appealed to cognate expressions in order to argue against degrees of consciousness. For example, Bayne, Hohwy, & Owen [2016: 407] say, "[T]he notion of degrees of consciousness is of dubious coherence […] [A] creature is conscious if and only if it possesses a subjective point of view. Arguably, the property of having a subjective point of view is not gradable—it cannot come in degrees." Similarly, Carruthers [2019] contends that "we can't make sense of degrees of phenomenal consciousness" (23) on the grounds that "[e]ither a mental state is like something for its subject to undergo, or it is not" (5) and that "[i]t is hard even to conceive of a case of a mental state that is partly like something to undergo, partly not" (5). And Tye [forthcoming: 8] criticizes the idea that some states are "literally […] more conscious" on the grounds that one state cannot "be an experience to a greater degree than another, as one person can be older or taller than another." To address the best version of this objection, we will consider a version that targets 'what it's like' expressions, though our arguments will generalize to the other cognate expressions.

In order for this objection to have any plausibility, it cannot be targeting straightforward substitutions of 'what it's like' in place of 'conscious'. This is because 'what it's like' isn't an adjective: it's a free relative

clause noun phrase.[23] Hence, such substitutions will often be unacceptable due to ungrammaticality, even for sentences without degree expressions (e.g., # 'A human is what it's like'). Instead, the objection must be about the acceptability of minimal variants of the target sentences that aim to preserve grammaticality:

(5*a*) There is more of what it's like to be a human than what it's like to be a fish.

(5*b*) If there's something that it's like to be a snail, then there's probably only a little bit.

(5*c*) AIs may soon be such that there's a little that it's like to be them.

(5*d*) Perhaps there is more that it's like to be in a psychedelic state than there is to be in a sober state.

(5*e*) If we stimulate your prefrontal cortex, then we hypothesize that there will become more of what it's like to be you than there was before.

Once we attempt to eliminate sources of ungrammaticality, it's not entirely obvious that these constructions remain unacceptable. Still, those who maintain that these sentences *are* unacceptable might take this—along with the dubious coherence of degrees of what-it's-likeness—as evidence against GRADABILITY.

But that reasoning is too quick. There are many expressions—including gradable adjectives—that are definable in terms of another expression without being replaceable with that expression in every sentence (even minimal variants that aim to preserve grammaticality). Here's an example: what it is for a door to be open is for it to have a non-zero degree of openness. But the acceptability of describing a door as being 'more open than the window' doesn't automatically entail the coherence of describing it as having 'more of a non-zero degree of openness than the window'. Plausibly,

[23] For a recent analysis of 'what it's like' expressions, see Stoljar [2016].

having a non-zero degree of openness doesn't come in degrees: an object either has a non-zero degree of openness or fails to have a non-zero degree of openness.

For the proponent of GRADABILITY, the situation is parallel for 'conscious'. To count as conscious simpliciter, an entity needs to have a non-zero degree of consciousness (a point we will argue for in the next section). This is plausibly equivalent to saying that there's something it's like to be that entity. If 'conscious' is used as a gradable adjective, then combining it with degree modifiers allows us to talk about the degree to which something is conscious. But having a non-zero degree of consciousness, or there being something it's like to be an entity, might not themselves be degreed properties: the entity either has a non-zero degree of consciousness or fails to have a non-zero degree of consciousness; there's either something it's like to be that entity, or there's nothing it's like to be that entity. If x has a higher degree of consciousness than y, then while this presumably requires x to have experiences that differ in character from y's, it doesn't have any bearing on the fact that both x and y have experiences.

Gradable adjectives can often be associated with another expression that denotes all and only those things that have a certain degree of the relevant property. But even though the gradable adjective itself can combine with degree expressions, it need not follow that the associated expression can. More generally, focusing on cognates of 'conscious' that might sound questionable in some constructions isn't obviously relevant to the question of whether 'conscious' is a gradable adjective. The link between copular phrases with adjectives (e.g., 'is conscious'), noun phrases ('consciousness / subjective point of view') and possessive phrases ('has consciousness / has a subjective point of view') is complex, and it's possible that they denote things of different ontological categories.[24] Analogous points apply to

[24] For example, Francez & Koontz-Garboden [2017] hold that positive form gradable adjectives ('is aware') describe individuals, associated mass nouns ('awareness') describe

progressive aspect forms ('having a subjective point of view' / 'being conscious').

These observations about the complex relationships between gradable adjectives and associated expressions undermine the 'what it's like' objection. It's widely accepted that for an entity to be conscious is for there to be something it's like to be that entity. But it doesn't follow that 'conscious' can always be acceptably replaced with a 'what it's like' expression or that we can coherently talk about degrees of what-it's-likeness.

§5 Minimality

Some opponents of GRADABILITY have assumed that if 'conscious' is gradable, then there must be borderline cases of consciousness. But—they contend—there aren't such borderline cases, so therefore 'conscious' isn't gradable.

This inference is fallacious: just because an adjective is gradable doesn't mean that it generates borderline cases. Gradable adjectives such as 'tall' and 'hot' generate borderline cases because they're *relative* gradable adjectives. But we will argue that 'conscious' isn't relative. Instead, there are both linguistic and philosophical reasons for thinking that when 'conscious' is used as a gradable adjective, it's used as a *minimal-standard absolute* gradable adjective.[25]

---

portions of qualities (where *qualities* are a sort of abstract mass entity), and associated possessive phrases ('has awareness') describe relations between an individual and an interval of a quality for which the individual bears the possessive relation to a portion in that interval. A variant of this approach that fits more naturally with a scale-based analysis of gradable adjectives would treat the associated mass nouns as describing scales and the possessive phrases as describing relations between an individual and a scale interval for which the individual is mapped to a degree in that interval (*ibid.*, 52). We are grateful to an anonymous reviewer for this point.

[25] Our point is that absolute gradable adjectives preclude one source of borderline cases: namely, borderline cases that arise due to contextually-variant thresholds associated with

§5.1 Relative vs. Absolute

Gradable adjectives can be either *absolute* or *relative*. The distinction turns on what degree of the associated scale marks the threshold for the positive form ('is tall', 'is flat', etc.) to truthfully apply.[26] For *absolute gradable adjectives*, the threshold degree is fixed, and is either always the minimum degree or always the maximum degree. For *relative gradable adjectives*, the threshold degree is context-dependent. In other words, each gradable adjective belongs to one of the following categories:

Minimal (absolute):	The threshold is the *minimum* degree on the scale.
Maximal (absolute):	The threshold is the *maximum* degree on the scale.
Relative:	The threshold is a non-minimum, non-maximum degree that varies across different contexts.

For example, 'tall' is relative: in one context, a woman might count as 'tall' because the threshold degree is the height of the average gymnast, while in another context, the same woman might fail to count as 'tall' because the threshold is the height of the average professional basketball player. By contrast, 'straight' is maximal and 'bent' is minimal: to count as strictly 'straight', an object must have the maximum degree of straightness,

---

the positive form. This is compatible with the view that there can be borderline cases of whether something counts as (say) 'open' due to imprecise uses of this adjective (see *fn*.27). Similarly, we don't mean to suggest that being a gradable adjective is necessary for generating borderline cases. Many predicates that allow for borderline cases aren't associated with a gradable adjective (for example, 'is a heap').

[26] More technically, sentences without explicit degree expressions are widely thought to include a *covert standard* that specifies the degree that the relevant individual is required to equal or exceed for the sentence to be true (see Kennedy & McNally 2005). For absolute gradable adjectives, the covert standard always specifies a fixed degree. For relative gradable adjectives, the covert standard is context-sensitive and specifies different degrees in different contexts.

whereas any amount of bentness is enough to make an object count as 'bent'.[27] Here are some other examples of adjectives that fall under each of the three categories:

Minimal	*Maximal*	*Relative*
'open'	'closed'	'tall'
'wet'	'dry'	'intelligent'
'bent'	'straight'	'hot'
'impure'	'pure'	'long'
'bumpy'	'flat'	'smart'

There are several tests used to classify gradable adjectives amongst these categories.[28] We will focus on three tests: entailment, comparison class, and modifier tests.

§5.2 The Entailment Test

Whether F is minimal, maximal, or relative predicts different entailments from '*x* is more F than *y*':[29]

[27] Some might worry that the thresholds for absolute gradable adjectives can also be context-dependent. For example, someone might describe a door as 'closed' even when there are a few millimeters between the door and the doorframe. But these *imprecise uses* of absolute gradable adjectives should be distinguished from relative gradable adjectives. On one view, imprecision is a pragmatic phenomenon: imprecise uses express false claims, but are felicitous in contexts with reduced expectations of precision (Kennedy & McNally 2005). On another view, imprecision is a semantic phenomenon: imprecise uses of gradable adjectives are interpreted relative to coarser grained scales, which allows them to express true claims (Sassoon & Zevakhina 2012). But on both views, the standard for absolute gradable adjectives remains a minimal or maximal degree of the relevant scale.

[28] See Kennedy [2007] for the original invocation of these tests.

[29] Sometimes the grammatically correct construction is 'F-er than'. For simplicity, we will formulate everything in terms of 'more F than'.

Minimal:	x is more F than y	$\Rightarrow$	x is F
		$\not\Rightarrow$	y is not F
Maximal:	x is more F than y	$\not\Rightarrow$	x is F
		$\Rightarrow$	y is not F
Relative:	x is more F than y	$\not\Rightarrow$	x is F
		$\not\Rightarrow$	y is not F

The reason the test works is as follows. If the comparative entails that x is F, then x's having a non-zero degree of F-ness is sufficient for 'F' to apply to x, meaning 'F' is minimal. If the comparative entails that y is not F, then y's failure to have the maximum degree of F-ness suffices for 'not F' to apply to y, meaning 'F' is maximal. If neither entailment holds, then x's having a non-zero degree of 'F'-ness isn't sufficient for 'F' to apply to x, and y's having a non-maximal degree of F-ness isn't sufficient for 'not F' to apply to y, meaning 'F' is relative. Consider some examples:

(6*a*) The front door is more open than the back door.
- $\Rightarrow$ The front door is open.
- $\not\Rightarrow$ The back door is not open.
- $\therefore$ 'open' is minimal.

(6*b*) The rod is straighter than the stick.
- $\not\Rightarrow$ The rod is straight.
- $\Rightarrow$ The stick is not straight.
- $\therefore$ 'straight' is maximal.

(6*c*) Aldo is taller than Bezawit.
- $\not\Rightarrow$ Bezawit is tall.
- $\not\Rightarrow$ Bezawit is not tall.
- $\therefore$ 'tall' is relative.

What about 'conscious'? Well, 'conscious' seems to exhibit the behavior of a minimal gradable adjective: if *x* is more conscious than *y*, then that seems to entail that *x* is conscious. For example:

(6*d*) The fish is more conscious than the snail.
⇒ The fish is conscious.
⇏ The snail is not conscious.
∴ 'conscious' is minimal.

Suppose that in this example 'conscious' is genuinely used as a gradable adjective (as opposed to, say, a metalinguistic comparative use). Then it seems to follow that the fish is conscious. It's hard to understand what it would mean to say ? 'The fish is more conscious than the snail, but (even) the fish isn't conscious'. But it seems coherent to say 'The fish is more conscious than the snail, but the snail is also conscious'. This pattern of entailments supports classifying 'conscious' as minimal.

§5.3 The Comparison Class Test

The comparison class test concerns whether a gradable adjective sounds acceptable with 'for'-phrases that specify a comparison class, such as expressions of the form '*x* is F for a G'. When such constructions are felicitous, the gradable adjective is likely relative. This is because the threshold for relative—but not absolute—gradable adjectives is affected by a comparison class. Compare:

(7*a*) She's tall for (a child / an American woman / a basketball player).
(7*b*) ? That's open for (a door / a secret cabinet / a snake's mouth).

Whereas (7*a*) sounds acceptable, (7*b*) sounds strange. Furthermore, the truth-conditions for (7*a*)—unlike (7*b*)—are affected by different choices of 'for'-phrase. What counts as 'tall' varies across contexts, but something

counts as 'open' if and only if it has a non-zero degree of openness, regardless of the context.

What about 'conscious'? Well, 'conscious' clearly patterns with absolute gradable adjectives here:[30]

(7*c*) ? Aldo is conscious for (a fish / a coma patient / a human).

That is, something counts as 'conscious' if and only if it has a non-zero degree of consciousness; the degree of consciousness that must be exceeded for a fish or coma patient to count as 'conscious' is always degree zero.

§5.4 The Modifier Test

The modifier test concerns a gradable adjective's acceptability—under its ordinary interpretation—with certain *endpoint-oriented degree modifiers*:[31]

[30] Brogaard [ms] also applies the comparison class test to 'conscious' and observes that it sounds unacceptable with the relevant type of 'for'-phrases. However, she concludes that 'conscious' belongs to a special class of relative gradable adjectives that the existing literature has overlooked and that fail the comparison class test. By contrast, we think the entailment and modifier tests—which Brogaard doesn't apply—are evidence that 'conscious' is a minimal gradable adjective. Our hypothesis explains why 'conscious' fails the comparison class test without requiring the endorsement of a new, special class of relative gradable adjectives.

[31] Strictly speaking, the modifier test provides information about scale structure, rather than direct information about the type of covert standard. That is, acceptability with lower endpoint-oriented modifiers requires the associated scale to have a minimum element, and acceptability with upper endpoint-oriented modifiers requires the associated scale to have a maximum element. It's widely accepted, however, that scale structure correlates with covert standard type (see Kennedy 2007): scales with a minimum element but no maximum element tend to be associated with minimal gradable adjectives, scales with a maximum element but no minimum element tend to be associated with maximal gradable adjectives, scales with both a maximum and a minimum element can be associated with either type of absolute gradable adjective, and scales with neither a minimum nor maximum element are always associated with relative gradable adjectives. Given this correlation, the modifier test provides indirect evidence about covert standard type.

Lower endpoint-oriented modifiers
'slightly', 'a little'

Upper endpoint-oriented modifiers
'completely', 'perfectly', 'totally', 'absolutely'

If a gradable adjective is acceptable with lower endpoint-oriented modifiers but not with upper endpoint-oriented ones, then that indicates that it's *minimal*. If a gradable adjective is acceptable with upper endpoint-oriented modifiers but not with lower endpoint-oriented ones, then that indicates that it's *maximal*. And if it's acceptable with neither group of endpoint-oriented modifiers, then that indicates that it's *relative*:

(8*a*) The dirt road is (slightly / a little / ? completely / ? perfectly / ? totally / ? absolutely) bumpy.

(8*b*) The dirt road is (? slightly / ? a little / completely / perfectly / totally / absolutely) flat.

(8*c*) Aldo is (? slightly / ? partly / ? a little / ? completely / ? perfectly / ? totally / ? absolutely) tall.

What about 'conscious'? From our experience, a good proportion of people's linguistic intuitions pattern as below:

(8*d*) (Humans / fish) are (slightly / a little / ? completely / ? perfectly / ? totally / ? absolutely) conscious.

On balance, the three tests—the comparison class, entailment, and modifier tests—best support classifying gradable uses of 'conscious' as minimal

standard.[32] In other words, having some non-zero degree of consciousness is always sufficient to count as 'conscious'. Those who favor classifying 'conscious' as maximal or relative must motivate different verdicts about the tests. Moreover, they also face some implausible consequences, which we will turn to next.

§5.5 Context-Sensitivity

Here's the first implausible consequence: if 'conscious' is relative, then whether an entity counts as 'conscious' depends on the context.

Recall that whether something counts as (say) 'tall' is context-dependent, unlike whether something counts as strictly 'straight' or 'bent'. A five-foot-six woman might count as 'tall' when we talk about her in one context but not when we talk about her in another context (say, depending on whether the relevant comparison class consists of gymnasts or basketball players). This is because the threshold degree for relative gradable adjectives is context-dependent, unlike for absolute gradable adjectives.

Is 'conscious' more like 'bent' or 'tall'? It's highly counterintuitive to hold that whether something counts as 'conscious' in the phenomenal sense depends on the context in which the adjective is used. For example, it's implausible that an ant could count as phenomenally 'conscious' when talking about it in one context but not when talking about it in another context (depending on whether the comparison class consists of, say, insects or land

[32] Some readers might find some of the results unclear. We suspect that this may be partly due to uncertainty about the associated scale for consciousness (even if we take it for granted that 'conscious' is gradable). For example, some views of degrees of consciousness entail that there's a maximal degree (such as a view that identifies degrees of consciousness with degrees of attention), while other views entail that there's no maximal degree (such as a view that identifies degrees of consciousness with amount of integrated information). If one is sympathetic to views that have implications for the structure of the consciousness scale, then that might influence one's reaction to the example sentences. Furthermore, the tests rely on linguistic intuitions, and our tacit knowledge of English might not provide clear verdicts for phenomenal, gradable uses of 'conscious' (see *fn*.8).

animals). Yet classifying gradable uses of 'conscious' as relative would entail this counterintuitive outcome. This is evidence that if 'conscious' is gradable, then it must be absolute rather than relative.

Moreover, while there's debate about which linguistic tests most reliably identify context-dependent expressions, some popular tests—such as tests suggested by Cappelen & Lepore [2003, 2004, 2005] and by Viebahn & Vetter [2016]—provide further evidence against classifying 'conscious' as such an expression.[33]

§5.6 The Phenomenal Differences Argument

Here's the second implausible consequence: if 'conscious' is either relative or maximal, then it follows that there can be two entities x and y such that x and y differ in their degree of consciousness, yet what it's like to be x doesn't differ from what it's like to be y.

To see why this is a consequence, suppose (per reductio) that (*a*) 'conscious' is either relative or maximal, that (*b*) x and y differ in their degree of consciousness, but that (*c*) neither x nor y surpasses the threshold for the positive form of 'conscious' to apply. This means one of the following must be true: (1) neither x nor y surpass the threshold (if 'conscious' is relative), or (2) neither x nor y has the maximal degree (if 'conscious' is

[33] Cappelen & Lepore [2003, 2005] argue that if a sentence *s* is context-sensitive, then there are some contexts where a speaker can truthfully utter a version of: 'There can be false utterances of '*s*' even though *s*'. For example: 'There can be false utterances of 'I am a philosopher' even though I am a philosopher'. This test arguably classifies 'conscious' as context-insensitive: it's hard to think of any context where there could be a true utterance of (say): 'There can be false utterances of 'slugs are conscious' even though slugs are conscious'. Viebahn & Vetter [2016: 8] argue that if an expression is context-sensitive, then it should be reasonably obvious to speakers that there's a high number of potential contents that can be expressed by that expression. For example, there are as many potential contents for 'I' as there are English speakers. This test also seems to classify 'conscious' as context-insensitive: it doesn't seem obvious that there's a high number of potential contents for 'conscious', at least not in the same way as 'I'.

maximal). Since neither entity surpasses the threshold (whether it's relative or maximal), neither entity counts as 'conscious'. Since, given the standard definition of 'phenomenally conscious', an entity is conscious just in case there's something it's like to be that entity, it follows that there's nothing it's like to be *x* and nothing it's like to be *y*. But if there's nothing it's like to be either *x* or *y*, then it trivially follows that what it's like to be *x* doesn't differ from what it's like to be *y*. Hence, taking 'conscious' to be either relative or maximal entails that there can be differences in degrees of consciousness without differences in what it's like.[34]

By contrast, such a scenario is impossible if 'conscious' is minimal. If *x* and *y* differ in their degree of consciousness, then either both have a positive degree of consciousness or one of them—say, *y*—has degree of consciousness zero.[35] If the former, then both will surpass the (minimal) threshold to count as 'conscious', so what it's like to be *x* will differ from what it's like to be *y* (since by hypothesis, they differ in degree of consciousness). If the latter, then there will be something it's like to be *x* but nothing it's like to be *y*, so once again what it's like to be *x* will differ from what it's like to be *y*. If 'conscious' is gradable, then only views that classify 'conscious' as minimal will be able to avoid the implausible consequence.[36]

[34] It's worth noting that someone who thinks 'conscious' is relative or maximal doesn't have to deny that *x* is conscious iff there's something it's like to be *x*. This is because they could hold that the positive form 'is conscious' applies iff *x* surpasses the threshold iff there's something it's like to be *x*. Our point, though, is that they'd have to deny the following principle: *x* and *y* differ in consciousness only if what it's like to be *x* differs from what it's like to be *y*. This result would violate Lee [2023]'s "difference criterion" on degrees of consciousness (if *x* and *y* differ in degree of consciousness, then what it's like to be *x* must differ from what it's like to be *y*).

[35] We assume there's no such thing as negative degrees of consciousness.

[36] Here's a challenge for the minimality classification: minimal gradable adjectives typically have maximal gradable adjectives as their antonyms, but it's not obvious that 'unconscious' or 'non-conscious' is gradable. However, there are other examples of apparently gradable adjectives with apparently non-gradable antonyms. As examples, 'possible',

We conclude that when 'conscious' is used as a gradable adjective, it's used as a minimal-standard absolute gradable adjective.

§6 Subjectivity vs. Phenomenal Character

We will end by turning to a more speculative hypothesis about why some consciousness researchers find gradable uses of 'conscious' infelicitous. We suspect that even after disentangling phenomenal consciousness from other senses of 'consciousness', there remain two ways of interpreting 'consciousness'.[37] Our argument appeals to a distinction developed in Lee [forthcoming] between what he calls 'subjectivity' and 'phenomenal character'. These two terms are characterized as follows:

subjectivity	$=_{\text{def}}$	what makes an entity feel some way at all.
phenomenal character	$=_{\text{def}}$	what it feels like to be an entity.

It's worth emphasizing that subjectivity is itself a phenomenal property (meaning a property that characterizes what it's like to have an experience). Therefore, any difference in subjectivity entails a difference in phenomenal character. While we will talk about different views of subjectivity via different candidates for the realizers of consciousness, subjectivity might be more accurately characterized as the phenomenal property that corresponds to those realizers.[38]

'awake', 'expensive', 'visible' and 'spotted' are widely classified as gradable (Lassiter [2010], Kennedy & McNally [2005: 359], Kennedy [2007]), but it arguably sounds marked to say '*x* is more (impossible / asleep / inexpensive / free / invisible / spotless) than *y*'.

[37] We will leave open whether 'phenomenal consciousness' is *polysemous* (meaning 'phenomenal consciousness' is associated with two distinct but related senses) vs. *indeterminate* (meaning 'phenomenal consciousness' has one sense that's compatible with expressing either of those properties). See Cruse [1986: 51] for more on ambiguity vs. indeterminacy.

[38] We will assume that any differences in the realizers—global workspace, higher-order thoughts, integrated information, etc.—entail different in phenomenology (and more specifically differences in subjectivity).

We won't explore the distinction between subjectivity and phenomenal character in detail here. Instead, we'll just briefly explain the distinction and offer some thoughts about its relevance for GRADABILITY. Note that the distinction isn't supposed to be motivated by ordinary language. Instead, it's a theoretical distinction that's motivated by two different ways of thinking about the meaning of 'phenomenal consciousness'.

To illustrate the distinction, Lee appeals to a common metaphor. It's often said that an entity is conscious just in case "the lights are on inside." But there are two interpretations of this metaphor. On the one hand, consciousness might be thought of as the inner light itself. On the other hand, consciousness might be thought of as the room that's illuminated (including the illumination itself). If we interpret 'consciousness' in the former way, then consciousness is subjectivity. If we interpret 'consciousness' in the latter way, then consciousness is phenomenal character.[39]

To make the distinction more concrete, consider *global workspace theory*, one of the most prominent contemporary theories of consciousness. According to the theory, an entity is conscious just in case it has a "global workspace," or a central executive system whose information is available to a wide range of other systems. But there are two ways of thinking about the meaning of 'consciousness', given global workspace theory. On the one hand, consciousness might be identified with the global workspace itself, since having a global workspace is what makes an entity feel some way at all. On the other hand, consciousness might be identified with the representations in the global workspace (alongside the global workspace itself), since the contents of those representations determine what it's like to be an

[39] Note that subjectivity and phenomenal character are defined as being intensionally equivalent: necessarily, *x* has subjectivity just in case *x* has phenomenal character. Because of this, standard tests for ambiguity are hard to apply to this case. For example, one standard test says that if 'F' is ambiguous, then there should be felicitous readings of '*x* is F, but *x* isn't F' (see Zwicky & Sadock 1975). But this test presupposes that some things are F in one sense but not in the other sense. However, it's impossible for something to have subjectivity without having phenomenal character (or vice versa).

entity. If we use 'consciousness' in the first way, then we're thinking of consciousness as subjectivity; if we use 'consciousness' in the second way, then we're thinking of consciousness as phenomenal character.

The distinction arises across a wide range of theories of consciousness. For example,

Most consciousness researchers don't explicitly draw this distinction; instead, it usually lies in the background of their theories. But from our experience, most recognize that this distinction is implicit in their theories once their attention is drawn to it. The following table lists some prominent theories of consciousness, as well as what those theories entail for subjectivity and for phenomenal character:

theory	**subjectivity**	**phenomenal character**
global workspace theory	global workspace	+ representations inside
higher-order theory	higher-order thoughts	+ first-order contents
integrated information theory	maximal ϕ-values	+ qualia structures
recurrent processing theory	recurrent feedback loops	+ contents

FIGURE 1: Theories of consciousness, subjectivity, and phenomenal character

We suspect that some researchers talking about consciousness have in mind subjectivity, some have in mind phenomenal character, and some have neither specifically in mind.[40] If this is right, then it may explain why some consciousness researchers find the target sentences infelicitous, even if GRADABILITY is true. This is because while there's a case to be made that subjectivity might come in degrees, it's harder to make sense of phenomenal character coming in degrees.

[40] 'Subjectivity' and 'phenomenal character' are intended to capture interpretations of 'consciousness' (a noun), rather than 'conscious' (an adjective). However, we think our hypothesis holds whether it's 'consciousness' or 'conscious' that's used: either way, researchers are talking and thinking about consciousness.

First, consider subjectivity. On many theories of consciousness, some creatures have more of whatever it is that makes an entity feel some way at all than other creatures. More technically, on some theories, some creatures can have a greater value than other creatures with respect to the property in virtue of which an entity feels some way at all.[41] For example, one entity might have greater attentional resources, or more higher-order thoughts, or a greater amount of integrated information, than another. Using the metaphor from earlier, it's easy to make sense of the idea that the light shines more brightly in some creatures than in others. To be clear, we take it to be an open question whether subjectivity does come in degrees. However, we suspect that when proponents of GRADABILITY use 'conscious' in gradable constructions, they generally have subjectivity in mind.

Second, consider phenomenal character. There's a reasonable case to be made that phenomenal character cannot come in degrees. One reason is that 'phenomenal character' is defined as what it's like to be an entity. But as discussed in §4, what it's like to be an entity is not itself a degreed property. Relatedly, Lee [forthcoming] suggests that 'phenomenal character' can also be characterized as the way it feels to be an entity. But it seems plausible that ways do not come in degrees either.[42]

We hypothesize that some who find the target sentences infelicitous may be thinking of consciousness as phenomenal character (rather than as subjectivity). Those focused on phenomenal character might find the sentences infelicitous because—as argued above—it's plausible that phenomenal character doesn't come in degrees. But we suspect that researchers who use 'conscious' as a gradable adjective are generally focused instead on subjectivity.

---

[41] See Lee [2023] for more detailed argument for this.

[42] Suppose—contra our arguments—that phenomenal character (and hence what-it's-likeness and ways of feeling) *can* come in degrees. That would undercut our hypothesis about why opponents of GRADABILITY find the target sentences infelicitous. But it would also strengthen our overall case for GRADABILITY: 'consciousness', however we interpret it, is something that can naturally be thought of as coming in degrees.

§7 The Semantics and the Metaphysics

Suppose we're right that 'conscious' sometimes functions as a gradable adjective. Does our conclusion have metaphysical consequences?

A natural (but mistaken) thought is that our conclusion implies that consciousness comes in degrees. If 'conscious' is sometimes used as a gradable adjective and gradable adjectives are associated with degreed properties, then it might seem to follow that there are degrees of consciousness. However, we think it's important to resist this quick route from the semantics to the metaphysics. On our view, the semantic conclusion that 'conscious' sometimes functions as a gradable adjective leaves open the metaphysical question of whether consciousness comes in degrees.

The relationship between the semantics of a term 'F' and the metaphysical nature of the property F depends on substantive semantic, metaphysical, and metasemantic issues. There are many views on which metaphysical conclusions (such as consciousness coming in degrees) don't automatically follow from semantic conclusions (such as 'conscious' functioning as a gradable adjective). This is because on many views, the meaning of an expression leaves open the metaphysical nature and the structure of things in reality.

As an example, here's a brief sketch of a picture that we find plausible. The picture appeals to *model-theoretic semantics*, where the meanings of expressions are relativized to models.[43] Each model includes a *domain*, which represents what exists according to the model, and an *interpretation function*, which assigns contents to expressions. An expression's content is an *intension*, which is a function from worlds to suitable extensions in the domain. On this view, dialects with semantically significant differences are interpreted relative to different models. The model for a dialect where 'conscious' is a gradable adjective will associate the adjective with a degreed property; the model for a dialect where 'conscious' is a non-gradable

[43] See Montague [1973], Partee [1975], and Barwise & Etchemendy [1989].

adjective will associate the adjective with a non-degreed property.[44] The extension of 'is conscious'—which will be a set of individuals according to both dialects—is predicted to be the same set of individuals for any given world (those that have a non-zero degree of consciousness for the first dialect, and those that have the non-degreed property of consciousness for the second).[45] But the question of which model best reflects the nature of reality is left open. In other words, the question of which model best reflects reality is a question for scientists and metaphysicians, rather than semanticists.

Do our semantic arguments have any metaphysical implications? We think that it's reasonable to draw a modest metaphysical conclusion: namely, that a certain form of argument against degrees of consciousness doesn't work. Suppose that 'conscious' never genuinely functions as a gradable adjective. Then, one might think, there are no semantic grounds for holding that consciousness comes in degrees. One might even think, in that case, that there *are* semantic grounds to *doubt* that consciousness comes in degrees. Our arguments undermine this line of reasoning: it's not the case that 'conscious' never functions as a gradable adjective. Therefore, while the semantics doesn't settle the metaphysics, the semantics undermines one path towards a metaphysical conclusion.

[44] More technically, for the first dialect the content assigned to 'conscious' yields extensions consisting of functions from individuals to degrees on a consciousness scale (at least according to a popular analysis of gradable adjectives; see Kennedy [2007]). For the second dialect, the content yields extensions consisting of functions from individuals to truth values.

[45] The positive form of a gradable adjective is generally thought to involve a covert element that yields an expression of the same semantic type as a non-gradable adjective (see *fn*.26), with an extension consisting of the set of individuals that have a degree of the property that exceeds the relevant threshold. We set aside the issue of how to identify the same individuals and worlds across the domains of distinct models.

Conclusion

The paper has revolved around GRADABILITY—the thesis that 'conscious' at least sometimes functions as a gradable adjective. We've argued that the balance of considerations favors GRADABILITY.

To set the stage, we argued that both proponents and opponents of GRADABILITY face some explanatory burdens: either we need an explanation for why some find gradable uses of 'conscious' infelicitous (if GRADABILITY is true), or an explanation for why some find gradable uses of 'conscious' felicitous (if GRADABILITY is false).

The most natural strategies for opponents of GRADABILITY strike us as insufficient for meeting this explanatory burden. It's implausible that all gradable uses of 'conscious' involve a non-phenomenal sense of 'conscious', and it's implausible that the most common kinds of non-standard uses of gradable adjectives can cover all gradable uses of 'conscious'. We've also argued that a common objection to GRADABILITY—namely, that 'what it's like' expressions cannot take on degree modifiers or feature in comparative constructions—doesn't work.

Over the course of the paper, we've offered several hypotheses for why some find gradable uses of 'conscious' infelicitous. We argued that some theorists may be overlooking the distinction between relative and absolute gradable adjectives, and conflating the question of whether consciousness comes in degrees with the question of whether there could be borderline cases of 'conscious'. We also hypothesized that some theorists may be interpreting 'consciousness' as expressing phenomenal character (rather than as subjectivity), which arguably cannot come in degrees.

At the start of the paper, we distinguished the metaphysical question of whether consciousness comes in degrees from the semantic question of whether 'conscious' is a gradable adjective. Our focus has been on the semantic question. On our view, the metaphysical question of whether consciousness comes in degrees is a question to be settled by philosophical analysis and scientific inquiry, rather than by semantics.

Nevertheless, we think our arguments concerning the semantic question can be deployed to advance the metaphysical question. There are some philosophers who have contended that there are semantic grounds for ruling out the possibility that consciousness comes in degrees. If it doesn't even make sense to talk about some creatures being more conscious than others, then perhaps the metaphysical question can be settled by careful attention to the semantics. There's rarely elaboration on what exactly these semantic arguments are, though. By investigating GRADABILITY, this paper has aimed to construct the best semantic arguments against degrees of consciousness. These include arguments appealing to the polysemy of 'conscious', to cognate expressions such as 'what it's like', and to connections between gradability and borderline cases. By dispelling those arguments, we've undermined one of the principal motivations for resisting degrees of consciousness.

Whether or not consciousness comes in degrees remains an open question. But—we've argued—there are no compelling semantic reasons for ruling out that thesis.[†]

[†] ACKNOWLEDGMENTS: We are grateful for helpful comments from Rob Stainton and an anonmyous referee for *Philosophy and Phenomenological Research*. This work was supported by the ERC Starting Grant project EVIL ("Expressing Value in Language," No. 101162444).

References

Barwise, J., Etchemendy, J. 1989. "Model-Theoretic Semantics", in *The Foundations of Cognitive Science*, Michael I. Posner (ed.). MIT Press.

Bayne, Tim ; Hohwy, Jakob & Owen, Adrian M. (2016). Are There Levels of Consciousness? *Trends in Cognitive Sciences* 20 (6):405-413.

Birch, Jonathan ; Schnell, Alexandra K. & Clayton, Nicola S. (2020). Dimensions of Animal Consciousness. *Trends in Cognitive Sciences* 24 (10):789-801.

Bostrom, N. (2023, April 12) What if A.I. sentience is a question of degree? *New York Times.* https://www.nytimes.com/2023/04/12/world/artificial-intelligence-nick-bostrom.html

Brogaard, B. (2010). *Degrees of consciousness*. Unpublished manuscript.

Burnett, H. (2017). *Gradability in natural language: Logical and grammatical foundations*. Oxford University Press.

Cappelen, H., & Lepore, E. (2003). Context Shifting Arguments. *Philosophical Perspectives*, *17*(1), 25–50. https://doi.org/10.1111/j.1520-8583.2003.00002.x

Cappelen, H., & Lepore, E. (2005). *Insensitive semantics: A defense of semantic minimalism and speech act pluralism*. Blackwell.

Carruthers, Peter (2019). *Human and Animal Minds: The Consciousness Questions Laid to Rest*. Oxford University Press.

Chalmers, David John (2022). *Reality+: Virtual Worlds and the Problems of Philosophy*. New York: W. W. Norton.

Condoravdi, C., & Gawron, J. M. (1996). The context-dependency of implicit arguments. In M. Kanazawa, C. Piñón, & H. de Swart (Eds.), *Quantifiers, Deduction, and Context* (pp. 1–32). CSLI Publications.

Cruse, D. A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.

Francez, I., & Koontz-Garboden, A. (2017). *Semantics and Morphosyntactic Variation: Qualities and the Grammar of Property Concepts*. Oxford: Oxford University Press.

Giannakidou, A., & Yoon, S. (2011). The subjective mode of comparison: Metalinguistic comparatives in Greek and Korean. *Natural Language & Linguistic Theory*, *29*(3), 621–655.

Gopnik A. Why babies are more conscious than we are. *Behavioral and Brain Sciences*. 2007;30(5-6):503-504. doi:10.1017/S0140525X0700283X

Hawthorne, J., & Logins, A. (2021). Graded epistemic justification. *Philosophical Studies*, *178*, 1845–1858.

Kaplan, D. (1989). Demonstratives. In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 481–563). Oxford University Press.

Kennedy, C., & McNally, L. (2005). Scale Structure, Degree Modification, and the Semantics of Gradable Predicates. *Language*, *81*.

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, *30*(1), 1–45. https://doi.org/10.1007/s10988-006-9008-0

Koch, Cristof (2011). *Consciousness: Confessions of a Romantic Reductionist*. MIT Press.

Landau, I. (2010). The explicit syntax of implicit arguments. *Linguistic Inquiry*, *41*(3), 357–388. https://doi.org/10.1162/LING_a_00001

Lasersohn, P. (1993). Lexical distributivity and implicit arguments. In U. Lahiri & A. Wyner (Eds.), *Proceedings of Semantics and Linguistic Theory (SALT)* (Vol. 3, pp. 145–161). Cornell University Linguistic Publications. https://doi.org/10.3765/salt.v3i0.2751

Lassiter, D. (2010). Gradable Epistemic Modals, Probability, and Scale Structure. In D. Lutz & N. Li (Eds.), *Proceedings of Semantics and Linguistic Theory (SALT)* (Vol. 20, pp. 197–215). CLC Publishings.

Lassiter, D. (2017). *Graded Modality: Qualitative and Quantitative Perspectives* (First edition). Oxford University Press.

Lee, Andrew Y. (2023). Degrees of Consciousness. *Noûs* 57 (3):553-575.

Lee, Andrew Y. (forthcoming). The Light & the Room. . In D. Curry & L. Daoust (Eds.), *Introducing Philosophy of Mind, Today.* Routledge.

Mankowitz, Poppy (2023). Not Half True. *Mind* 132 (525):84-112.

Montague, R. (1973). The proper treatment of quantification in ordinary English. In Patrick Suppes, Julius Moravcsik & Jaakko Hintikka (eds.), *Approaches to Natural Language*. Dordrecht. pp. 221--242.

Morzycki, M. (2011). Metalinguistic comparison in an alternative semantics for imprecision. *Natural Language Semantics*, *19*(1), 39–86. https://doi.org/10.1007/s11050-010-9063-5

Mckilliam, Andy Kenneth (2020). What is a global state of consciousness? *Philosophy and the Mind Sciences* 1 (II).

Partee, B. "Montague Grammar and Transformational Grammar." *Linguistic Inquiry*, vol. 6, no. 2, 1975, pp. 203–300.

Pavese, Carlotta (2017). Know-How and Gradability. Philosophical Review 126 (3):345-383.

Sassoon, G. W., & Zevakhina, N. (2012). Granularity shifting: Experimental evidence from degree modifiers. In A. Chereches (Ed.), *Proceedings of Semantics and Linguistic Theory (SALT)* (Vol. 22, pp. 226–246). Cornell University Linguistic Publications.

Silva Jr, Paul & Siscoe, Robert Weston (forthcoming). Awareness By Degree. Philosophy and Phenomenological Research.

Siscoe, R. W. (2022). Rational supererogation and epistemic permissivism. *Philosophical Studies*, *179*(2), 571–591.

Stoljar, Daniel (2016). The Semantics of 'What it's like' and the Nature of Consciousness. *Mind* 125 (500):1161-1198.

Sutskever, I, Twitter 2022 [@ilyasut]. (n.d.). *Tweets* [Ilya Sutskever]. Retrieved February 28th, 2024, from https://twitter.com/ilyasut/status/1491554478243258368?lang=en

Tononi, G. An information integration theory of consciousness. *BMC Neurosci* **5**, 42 (2004).

Tye, Michael (forthcoming). How I learned to stop worrying and love panpsychism. Journal of Consciousness Studies.

Viebahn, E., & Vetter, B. (2016). How many meanings for 'may'? The case for modal polysemy. *Philosophers' Imprint*, *16*(10), 1–26.

Zwicky, A. M., & Sadock, J. M. (1975). Ambiguity tests and how to fail them. In J. P. Kimball (Ed.), *Syntax and Semantics* (Vol. 4, pp. 1–36). Academic Press.