

Phenomenal Consciousness, Defined and Defended as Innocently as I Can Manage

Commentary on Keith Frankish, “Illusionism as a Theory of Consciousness”

Eric Schwitzgebel
Department of Philosophy
University of California at Riverside
Riverside, CA 92521-0201
USA

Abstract:

Phenomenal consciousness can be conceptualized innocently enough that its existence should be accepted even by philosophers who wish to avoid dubious epistemic and metaphysical commitments such as dualism, infallibilism, privacy, inexplicability, or intrinsic simplicity. Definition by example allows us this innocence. Positive examples include sensory experiences, imagery experiences, vivid emotions, and dreams. Negative examples include growth hormone release, dispositional knowledge, standing intentions, and sensory reactivity to masked visual displays. Phenomenal consciousness is the most folk psychologically obvious thing or feature that the positive examples possess and that the negative examples lack, and which preserves our ability to wonder, at least temporarily, about antecedently unclear issues such as consciousness without attention and consciousness in simpler animals. As long as this concept is not empty, or broken, or a hodgepodge, we can be phenomenal realists without committing to dubious philosophical positions.

1. Introduction.

Keith Frankish argues that phenomenal consciousness does not really exist. I, along with most other Anglophone philosophers who have written on the issue, think that phenomenal consciousness does exist.

Frankish can be interpreted as posing a dilemma for defenders of the real existence of phenomenal consciousness – *phenomenal realists*, as I will call them. On Horn 1 we find inflated views about what phenomenal consciousness involves: infallibility, or metaphysical dualism, or some other dubious philosophical commitments. If phenomenal consciousness requires any of *those* features, it probably doesn't exist. On Horn 2 we find views so deflationary as to be tantamount to the non-existence of the originally intended phenomenon. Frankish argues that we must choose between thinking of phenomenal consciousness in so inflated a way that there fails to be any such thing or so deflationary a way that it has effectively vanished into something else (e.g., dispositions to make certain sorts of judgments).

The best way to meet Frankish's challenge is to provide something that the field of consciousness studies in any case needs: a clear definition of phenomenal consciousness, a definition that targets a phenomenon that is both *substantively interesting* in the way that phenomenal consciousness is widely thought to be interesting but also *innocent of problematic metaphysical and epistemological assumptions*. In Section 2, I will attempt to do this.

One necessary condition on being substantively interesting in the relevant sense is that phenomenal consciousness should retain at least a superficial air of mystery and epistemic difficulty, rather than collapsing immediately into something as straightforwardly deflationary as dispositions to verbal report, or functional "access consciousness" in Block's (1995/2007) sense,

or an “easy problem” in Chalmers’ (1995) sense. If the reduction of phenomenal consciousness to something physical or functional or “easy” is possible, it should take some work. It should not be *obviously* so, just on the surface of the definition. We should be able to wonder how consciousness could possibly arise from functional mechanisms and matter in motion. Call this the *wonderfulness* condition.

2. Defining Consciousness by Example.

Unfortunately, the three most obvious, and seemingly respectable, approaches to definition all fail. Phenomenal consciousness cannot be defined *analytically*, in terms of component concepts (as “rectangle” might be defined as a right angled-planar quadrilateral). It is a foundationally simple concept, not divisible into component concepts. Even if it were to prove upon sufficient reflection to be analytically decomposable without remainder, *defining* it in terms of some hypothesized decomposition, right now, at our current stage of inquiry, would beg the question against researchers who would reject such a decomposition. Widespread disagreement also makes it inadvisable to define phenomenal consciousness *functionally*, in terms of the causal role it normally plays (as “heart” might be defined as the organ that normally plays the causal role of pumping blood). It’s too contentious what causal role, if any, phenomenal consciousness might have. Nor, for present purposes, can phenomenal consciousness be adequately defined by *synonymy*, since Frankish’s inflation-or-deflation dilemma applies equally to all of the nearby terms and phrases like “qualia”, “what-it’s-like-ness”, or “stream of experience”.

The best approach is *definition by example*. Definition by example can sometimes work well, if one provides diverse positive and negative examples and if the target concept is natural

enough that the target audience can be trusted to latch onto that concept once sufficient positive and negative examples are provided. I might say “by *furniture* I mean tables, chairs, desks, lamps, ottomans, and that sort of thing; and not pictures, doors, sinks, toys, or vacuum cleaners”. Hopefully, you will latch on to approximately the relevant concept (e.g., not being tempted to think of a ballpoint pen as furniture but being inclined to think that a dresser probably is). I might even define *rectangle* by example, by sketching out for you a variety of instances and nearby counterinstances (triangles, parallelograms, trapezoids, open-sided near-rectangles). Hopefully, you get the idea.

Definition by example is a common approach among recent phenomenal realists. I interpret Searle (1992, p. 83), Block (1995/2007, p. 166-168), and Chalmers (1996, p. 4) as aiming to define phenomenal consciousness by a mix of synonymy and appeal to example, plus maybe some version of the wonderfulness condition. All three attempts are, in my view, reasonably successful. However, all three attempts also have three shortcomings, which I aim to repair here. First, they are not sufficiently clear that they *are* definitions by example, and consequently they don’t sufficiently invite the reader to reflect on the conditions necessary for definition by example to succeed. Second, perhaps partly as a result of the first shortcoming, they don’t provide enough of the negative examples that are normally part of a good definition by example. Third, they are either vague about the positive examples or include needlessly contentious cases. Siewert (1998, ch. 3) is somewhat clearer on these points, but still limited in his range of negative examples and in his exploration of the conditions of failure of definition by example.

I want to highlight one crucial background condition that is necessary for definition by example to succeed. There must be an *obvious* or *natural* category or concept that the audience

will latch onto once sufficiently many positive and negative examples have been provided. In defining *rectangle* by example for my eight-year-old daughter, I might draw all of the examples with a blue pen, placing the positive examples on the left and the negative examples on the right. In principle, she might leap to the idea that “rectangle” refers to retangularly-shaped-things-on-the-left, or she might be confused about whether red figures can also be rectangles, or she might think I am referring to spots on the envelope rather than to the drawn figures. But that’s not how the typical enculturated eight-year-old human mind works. The definition succeeds because I know she’ll latch onto the intended concept, rather than some less obvious concept that fits the cases.

Defining *phenomenal consciousness* by example requires that there be *only one* obvious or readily adopted concept or category that fits with the offered examples. I do think that there is probably only one obvious or readily adopted category in the vicinity, at least once we do some explicit narrowing of possible candidates. In Section 3, I will discuss concerns about this assumption.

Let’s begin with positive examples. The word “experience” is sometimes used non-phenomenally (e.g., “I have twenty years of teaching experience”). However, in normal English it often refers to phenomenal consciousness. Similarly for the adjective “conscious”. I will use those terms in that way now, hoping that when you read them they will help you latch onto relevant examples of phenomenal consciousness. However, I will not always rely on those terms. They are intended as aids to point you toward the examples rather than as (possibly circular or synonymous) components of the definition.

Sensory and somatic experiences. If you aren’t blind and you think about your visual experience, you will probably find that you are having some visual experience right now. Maybe

you are visually experiencing black text on a white page. Maybe you are visually experiencing a computer screen. If you press the heels of your palms firmly against your closed eyes for several seconds, you will probably notice a swarm of bright colors and figures, called phosphenes. All of these visual goings-on are examples of phenomenal consciousness. Similarly, you probably have auditory experiences if you aren't deaf – at least when you stop to think about it. Maybe you hear the hum of your computer fan. Maybe you hear someone talking down the hall. In a sufficiently quiet environment, you might even hear the rush of blood in your ears. If you cup one hand silently over one ear, you will probably notice the change in ambient sound. If you stroke your chin with a finger, you will probably have tactile experience. Maybe you are feeling the pain of a headache. If you sip a drink right now, you will probably experience the taste and feel of the drink in your mouth. If you close your eyes and think about where your limbs are positioned, you might have proprioceptive experience of your bodily posture, which you might notice becoming vaguer if you remain motionless for an extended period.

Conscious imagery. Maybe there's unconscious imagery, but if there is, it's doubtful that you will be able to reflect upon an instance of it at will. Try to conjure a visual image – of the Eiffel Tower, say. Try to conjure an auditory image – of the tune of “Happy Birthday”, for example. Imagine it sung in your head. Try to conjure a motor image. Imagine how it would feel to stretch your arms back and wiggle your fingers. You might not succeed in all of these imagery tasks, but hopefully you succeeded in at least one, which you can now think of as another example of phenomenal consciousness.

Emotional experience. Presumably, you have had an experience of sudden fear on the road, during or after a near-accident. Presumably, you have felt joy, surprise, anger, disappointment, in various forms. Maybe there is no unified core feeling of “fear” or “joy” that

is the same from instance to instance. No matter. Maybe all there is to emotional experience is various sorts of somatic, sensory, and imagery experiences. That doesn't matter either. Think of some occasions on which you have vividly felt what you would call an emotion. Add those to your list of examples of phenomenal consciousness.

Thinking and desiring. Probably you've thought to yourself something like "what a jerk!" when someone has behaved rudely to you. Probably you've found yourself craving a dessert. Probably you've stopped to try to plan out, deliberately in advance, the best route to the far side of town. Probably you've found yourself wishing that Wonderful Person X would notice and admire you. Presumably not all of our thinking and desiring is phenomenally conscious in the intended sense, but presumably any instances you can now vividly remember or create are or were phenomenally conscious in the intended sense. Add these to your stock of positive examples. Again, it doesn't matter if these experiences aren't clearly differentiated from other types of experience that we've already discussed.

Dream experiences. Although in one sense of "conscious" we are not conscious when we dream, according to both mainstream scientific psychology and the folk understanding of dreams, dreams are phenomenally conscious – involving sensory or quasi-sensory experience, or maybe instead only imagery, and often some emotional or quasi-emotional component, like dread of the monster who is chasing you.

Other people. Bracketing radical skepticism about other minds, we normally assume that other people also have sensory experiences, imagery, emotional experiences, conscious thoughts and desires, and dreams. Count these, too, among the positive examples.

Negative examples. Not everything going on inside of your body is part of your phenomenal consciousness. You do not, presumably, have phenomenally conscious experience

of the growth of your fingernails, or of the absorption of lipids in your intestines, or of the release of growth hormones in your brain – nor do other people experience such things in themselves. Nor is everything that we normally classify as mental part of phenomenal consciousness. Before reading this sentence, you probably had no phenomenal consciousness of your disposition to answer “twenty-four” when asked “six times four”. You probably had no phenomenal consciousness of your standing intention to stop for lunch at 11:45. You presumably have no phenomenal consciousness of the structures of very early auditory processing. If a visual display is presented for several milliseconds and then quickly masked, you do not have visual experience of that display (even if it later influences your behavior). Nor do you have sensory experience of every aspect of what you know to be your immediate environment: no visual experience of the world behind your head, no tactile experience of the smooth surface of your desk that you can see but aren’t presently touching. Nor do you have pain experience, presumably, in regions outside your body, nor do you literally experience other people’s thoughts and images. We normally think that dreamless sleep involves a complete absence of phenomenal consciousness.

Phenomenal consciousness is the most folk psychologically obvious thing or feature that the positive examples possess and that the negative examples lack. I do think that there is one very obvious feature that ties together sensory experiences, imagery experiences, emotional experiences, dream experiences, and conscious thoughts and desires. They’re all *conscious experiences*. None of the other stuff is experienced (lipid absorption, the tactile smoothness of your desk, etc.). I hope it feels to you like I have belabored an obvious point. Indeed, my argumentative strategy relies upon this obviousness.

You must not try to be too clever and creative here! Of course you could invent a *new and non-obvious* concept that fits with the examples. You could invent some quus-like feature or “Cambridge property” like *being conscious and within 30 miles of Earth’s surface* or *being referred to in a certain way by Eric Schwitzgebel in this essay*. Or you could pick out some scientifically constructed but folk-psychologically non-obvious feature like accessibility to the “central workspace” or in-principle-reportability-by-a-certain-type-of-cognitive-mechanism. Or you could pick out a feature of the sort Frankish suggests, like “quasi-phenomenality” or presence of the disposition to *judge* that one is having wonderful conscious experiences. None of those are the feature I mean. I mean the *obvious* feature, the thing that kind of smacks you in the face when you think about the cases. That one!

Don’t try to analyze it yet. Do you have an analysis of “furniture”? I doubt it. Still, when I talk about “furniture” you know what I’m talking about and you can sort positive and negative examples pretty well, with some borderline cases. Do the same with phenomenal consciousness. Even you can do this, Keith! Let yourself fall into it. Save the analysis, reduction, and metaphysics for later.

Maybe scientific inquiry and philosophical reflection will reveal all the examples to have some set of functional properties in common, or to be reducible to certain sorts of brain processes, or whatever. That’s fine – a variety of unifying features can also be found for “square” and probably “furniture”. This doesn’t prevent us from defining such terms by example while remaining open-minded and non-commisive about theoretical questions that might be answered in later inquiry.

3. Contentious Cases and Wonderfulness.

Some consciousness researchers think that phenomenal consciousness is possible without attention – for example, that you are constantly phenomenally conscious of the feeling of your feet in your shoes even though you rarely attend to your feet or shoes. Others think consciousness is limited only to what is in attention. Some consciousness researchers think that phenomenal consciousness is exhausted by sensory, imagery, and emotional experiences, while others think that phenomenal consciousness comes in a wider range of uniquely irreducible kinds, possibly including imageless thoughts, an irreducible sense of self, or feelings of agency.

I have avoided committing on these issues by restricting the examples in Section 2 to what I think are likely to be uncontentious cases. I did not, for example, list a peripheral experience of the feeling of your feet in your shoes among the positive examples, nor did I list a nonconscious knowledge of the state of your feet among the negative examples. This leaves open the possibility that there are two or more fairly natural concepts that fit with the positive and negative examples and differ in whether they include or exclude such contentious cases. For example, if phenomenal consciousness substantially outruns attention, both the intended concept of *phenomenal consciousness* and the narrower concept of *phenomenal-consciousness-along-with-attention* adequately match the positive and negative examples.

Similarly, consciousness might or might not always involve some kind of reflective self-knowledge, some awareness of oneself *as* conscious. I intend the concept as initially open on this question, prior to careful introspective and other evidence.

You might find it introspectively compelling that your own stream of phenomenally conscious experience does, or does not, involve constant experience of your feet in your shoes, or reflective self-knowledge, or an irreducible sense of agency. Such confidence is, in my view, often misplaced (Schwitzgebel 2011). But regardless of whether such confidence is misplaced,

the intended concept of *phenomenal consciousness* does not build in, *as a matter of definition*, that consciousness is limited (or not) to what's in attention, or that it includes (or fails to include) phenomena such as an irreducible awareness of oneself as an experiencing subject. If it seems to you that there are two equally obvious concepts here, one of which is definitionally commissive on such contentious matters and another of which leaves such questions open to introspective and other types of evidence, my intended concept is the less commissive one. This is in any case probably the more obvious concept. We can *argue* about whether consciousness outruns attention; it's not normally antecedently stipulated.

It is likewise contentious what sorts of organisms are phenomenally conscious. Do snails, for example, have streams of phenomenally conscious experience? If I touch my finger to a snail's eyestalk, does the snail have visual or tactile phenomenology? If *phenomenal consciousness* meant "sensory sensitivity" we would have to say yes. If *phenomenal consciousness* meant "processes reportable via a cognitively sophisticated faculty of introspection", we would have to say no. I intend neither of these concepts, but rather a concept that doesn't settle the question as a straightforward matter of definition – and again I think this is probably the more typical concept to latch onto in any case.

It is this openness in the concept that enables it to meet the wonderfulness condition I introduced at the end of Section 1. One can *wonder* about the relationship between phenomenal consciousness and reportability, wonder about the relationship between phenomenal consciousness and sensory sensitivity, wonder about the relationship between phenomenal consciousness and any particular functional or biological process. One can wonder whether your stream of phenomenal consciousness could survive your bodily death. Maybe a bit of investigation will definitively settle these questions. Wonder doesn't have to be permanent.

Wonder is compatible even with the demonstrable mathematical impossibility of some of the epistemically open options: Before doing the calculation, one can wonder if and where the equation $y = x^2 - 2x + 2$ crosses the x axis. The “wonderfulness” condition as I intend it here does not require any kind of insurmountable “epistemic gap” – only a moment’s epistemic breathing space.

I suggest that there is one folk psychologically obvious concept, perhaps blurry-edged, that fits the positive and negative examples while leaving the contentious examples open and permitting wonder of the intended sort. That’s the concept of phenomenal consciousness.

4. Problematic Assumptions?

Back to Frankish’s dilemma. We get poked by Horn 1 if we commit to anything metaphysically or epistemically dubious in committing to the existence of phenomenal consciousness. We get poked by – or rather, warmly invited to – Horn 2 if we end up with so deflationary a concept of phenomenal consciousness that it ends up just being some “easy” straightforwardly functional or physical concept.

Frankish offers a nice list of dubious commitments that I agree it would be good not to build into the definition of consciousness. Let me now disavow all such commitments – consistently, I hope, with everything I have written so far. Phenomenally conscious experiences need not be simple, nor ineffable, nor intrinsic, nor private, nor immediately apprehended. They need not have non-physical properties, be inaccessible to third-person science, or be inexplicable in physical terms. My definition by example did not, I believe, commit me on any such questions. My best guess is that all of those claims are false, if intended as universal generalizations about phenomenal consciousness. (However, if some such feature turns out to be

present in all of the examples and thus, by virtue of its presence, in some sense indirectly or implicitly “built into” the definition by example, so be it. Such indirect commitments shouldn’t implausibly inflate our target. The troubling commitments are the inflationary ones which are *absent* from the target mental states.)

My definition did commit me to a fairly strong claim about folk psychology: that there is a single obvious folk-psychological concept or category that matches the positive and negative examples. But that’s a rather different sort of commitment.

I also committed to realism about that concept or category: The folk category is not empty or broken but rather picks out a feature that (most of) the positive examples share and the negative examples presumably lack. If the target examples had nothing important in common and were only a hodgepodge, this assumption would be violated. This is a substantive commitment, but not a dubious one I hope. (However, if the putative negative examples failed to be negative, as in some versions of panpsychism, we might still be able to salvage the concept, by targeting the feature that the positive examples have and that the negative examples are *falsely assumed* to lack.)

The wonderfulness condition involves a mild epistemic commitment in the *neighborhood* of non-physicality or non-reducibility. The wonderfulness condition commits to its being not straightforwardly obvious as a matter of definition what the relationship is between phenomenal consciousness and cognitive functional or physical processes. This commitment is quite compatible with the view that a clever a priori or empirical argument could someday show, perhaps even has already shown, that phenomenal consciousness is reducible to or identical to something functional or physical.

Frankish's quasi-phenomenality, characterized in terms of our dispositions to make phenomenal judgments, does not appear to meet the wonderfulness condition. (See also Frankish 2012's "zero qualia".) It does not leave open, even for a moment, the question of whether phenomenal consciousness might be present even in the absence of a certain cognitive functional feature: the disposition to make phenomenal judgments. I do think that question is open, at least for a moment – and probably for much more than a moment. I wonder, for example, whether snails might be conscious despite (presumably) their not being disposed to reach phenomenal judgments about their experience. I wonder whether we might have fleeting, unattended conscious experiences even if we are not disposed to reach judgments about them. I even wonder whether group entities like the United States might possess phenomenal consciousness at a group level, despite (presumably) no tendency to judge that they are doing so (though I doubt many people will join me in wondering about this).

After being invited to consider the positive and negative examples, someone might say, "I'm not sure I understand. What *exactly* do you mean by phenomenal consciousness?" At this point, it is tempting to clarify by making some epistemic or metaphysical commitments – whatever commitments seem plausible to you. You might say, "those events with which we are most directly and infallibly acquainted" or "the kinds of properties that can't be reduced to physical or functional role". Please don't! Or at least, don't build these commitments into the definition. Such commitments risk introducing doubt or confusion in people who aren't sure they accept such commitments. Maybe it's okay to say, "that about which it has often been *believed* we have direct, infallible access and *believed* to be irreducible to the physical". But let the examples do the work.

Here's a comparison: You are trying to teach someone the concept "pink". Maybe her native language doesn't have a corresponding term (as we don't have a widely used term for pale green). You have shown her a wide range of pink things (a pink pen, a pink light source, a pink shirt, pictures and photos with various shades of pink in various natural contexts); you've verbally referenced some famously pink things such as cherry blossoms and ham; you've shown her some non-pink things as negative examples (medium reds, pale blues, oranges, etc.). It would be odd for her to ask, "so do you mean this-shade-and-mentioned-by-you?" or "must 'pink' things be less than six miles wide?" It would be odd for her to insist that you provide an analysis of the metaphysics of pink before she accepts it as a workable concept. You might be open about the metaphysics of pink. It might be helpful to point, noncommittally, to what some people have said ("well, some people think of pink as a reflectance property of physical objects"). But lean on the examples. If she's not colorblind, and not perverse, there's something obvious that the positive instances share, which the negative examples lack, which normal people will naturally latch onto well enough, if they don't try too hard to be creative or insist on an analysis first, and if you don't confuse things by introducing dubious theses: They're all *pink*. This is a perfectly good way to teach someone the concept *pink*, well enough that she can confidently affirm that pink things exist (perhaps feeling baffled how anyone could deny it), sorting future positive and negative examples in more or less the consensus way, except perhaps in borderline cases (e.g., near-red) and contentious cases (e.g., someone's briefly glimpsed socks). My view is that the concept of *phenomenal consciousness* can be approached in the same manner.

I want, and I think we can reasonably have, and I think the most natural understanding of "consciousness" already gives us, room to wonder about certain things. We needn't commit

straightaway to either a reductionist picture on which everything is physical stuff, entirely mundane, or to what Frankish calls a “radical realist” picture on which consciousness somehow transcends the physical. If I had to bet, I’d bet on the mundane, but I don’t want to build it right into my conceptualization of consciousness. I want as innocent a concept as I can manage, which leaves the possibilities epistemically open.¹

References:

- Block, N. (1995/2007) On a confusion about a function of consciousness, in N. Block, *Consciousness, function, and representation*, Cambridge, MA: MIT.
- Chalmers, D.J. (1995) Facing up to the problem of consciousness, *Journal of Consciousness Studies*, **2** (3), 200-219.
- Chalmers, D.J. (1996) *The conscious mind*, Oxford: Oxford.
- Frankish, K. (2012) Quining diet qualia. *Consciousness & Cognition*, **21** 667-676.
- Schwitzgebel, E. (2011) *Perplexities of consciousness*, Cambridge, MA: MIT.
- Searle, J.R. (1992) *The rediscovery of the mind*, Cambridge, MA: MIT.
- Siewert, C. (1998). *The significance of consciousness*. Princeton: Princeton.

¹ For helpful discussion, thanks to Keith Frankish, Pauline Price, and commenters on my related blog post at the Splintered Mind.