Experiencers and the Ambiguity Objection[1]

Justin Sytsma

It is often asserted that we should believe that phenomenal consciousness exists because it is pretheoretically obvious. If this is the case, then we should expect lay people to categorize mental states in roughly the way that philosophers do, treating prototypical examples of (supposed) phenomenally conscious mental states similarly. Sytsma and Machery (2010) present preliminary evidence that this is not the case. They found that participants happily ascribed seeing red to a simple robot but denied that the robot felt pain. The most prominent response to this work has been the *ambiguity objection*, which charges that participants were interpreting ascriptions of seeing red in a purely informational way, such that their attributions of "seeing red" to the robot do not speak to the question of whether they recognize the phenomenality of this state. Peressini (2014) pushes an especially interesting version of the objection, presenting new empirical evidence and suggesting that lay people do in fact have a concept of phenomenality. In this paper, I respond to Peressini's objections, and the ambiguity objection more generally, arguing that the new data does not undermine Sytsma and Machery's conclusion.

Explaining phenomenal consciousness is one of the central concerns in contemporary philosophy of mind. Many hold both that phenomenal consciousness exists and that its existence is scientifically puzzling. It is claimed that unlike other phenomena that have resisted scientific explanation, phenomenal consciousness poses a distinctively hard problem in that we don't know how to go about offering a scientific account of it in either outline or detail. Looked at from the reverse angle, however, the seeming scientific intractability of the supposed phenomenon motivates skepticism. It provides reason to ask why we should believe that phenomenal consciousness exists in the first place.

The standard justification for belief in phenomenal consciousness in the literature is to simply claim that it is pretheoretically obvious, that it is "the most central and manifest aspect of

[1] I want to thank Eyuphan Özdemir for helpful comments on an earlier draft.

our mental lives" (Chalmers 1995, 207). But if this is true, then we should expect that lay people—people without training in philosophy or consciousness studies—will conceptualize our mental lives in roughly the same way that philosophers do, recognizing something important in common between prototypical examples of phenomenally conscious mental states. Across a number of articles, however, I've presented evidence that by and large lay people do not attribute mental states in a way that indicates such recognition.[2]

Most notably, in Sytsma and Machery (2010; SM) we offer empirical evidence that lay people tend to treat two prototypical examples of (supposed) phenomenally conscious mental states dissimilarly, attributing seeing red to a simple robot but denying that the robot feels pain. In contrast, philosophers tend to deny both attributions. From this, we argue that lay people do not categorize mental states in the way that philosophers do, and specifically that they do not categorize mental states like seeing red and feeling pain on the basis of recognizing that they have something important in common—that they are both phenomenal states. In short hand, we argue that lay people by and large lack the philosophical concept of phenomenal consciousness.

Anthony Peressini (2014) offers an interesting and nuanced objection to the work of SM, giving a version of what I have termed the *ambiguity objection* (Sytsma 2016). In general, the ambiguity objection contends that lay people possess something like the philosophical concept of phenomenal consciousness, but that they do not apply it uniformly in SM's cases. The claim is that unlike philosophers, lay people read "see red" in an informational sense akin to "detect red." As such, it is argued that lay participants' willingness to attribute seeing red to a simple robot does not indicate anything with regard to whether or not they have a concept of phenomenal

[2] See Sytsma 2009, 2010a, 2012, 2014a; Sytsma and Machery 2010, 2012; Sytsma and Özdemir ms1, ms2. This work falls under the area of experimental philosophy of mind. For a brief introduction focused on mental state attributions, see Machery and Sytsma 2011. For broader introductions to work in experimental philosophy of mind, see Part IIC of Sytsma and Buckwalter 2016 and Sytsma 2014b.

consciousness, but simply reflects the way that they read the question. We offered a number of responses to this objection in our original article, starting with the charge that absent a motivating explanation for this difference in readings between philosophers and lay people, the objection is ad hoc. Peressini responds by offering a further explanation. He provides evidence suggesting that there is a lay concept of "experiencers" that categorically divides people from robots. Peressini then argues that people think a phenomenal reading only applies to experiencers, and so instead adopt an informational reading in attributing seeing red to the robot.

In this paper, I respond to Peressini's version of the ambiguity objection and consider whether the results of his empirical study undermine SM's conclusion. I begin in Section 1 by describing the primary evidence that SM provide. In Section 2, I articulate the ambiguity objection and detail SM's responses to it. In Section 3, I turn to Peressini's counter-response, arguing that it does not make a compelling case in favor of the ambiguity objection. Finally, I consider Peressini's empirical work, discussing the first part of his study in Section 4 and the second part of his study in Section 5.

1. The Ambiguity Objection

SM argue that lay people by and large do *not* employ the philosophical concept of phenomenal consciousness, or something suitably similar, in making mental state ascriptions. And we argue that this undermines a standard justification given for believing in phenomenal consciousness in the first place—that it is pretheoretically obvious. We reason that if this justification is accurate, then we would expect to find that lay people classify mental states in roughly the same way that philosophers do, treating prototypical examples of phenomenally conscious mental states similarly. But the evidence suggests that they don't.

In our article, we note that there is disagreement among philosophers on what is meant by "phenomenal consciousness," but present what we take to be the standard understanding of the phrase. On this understanding, an entity is phenomenally conscious just in case it has phenomenally conscious mental states, and a mental state is phenomenally conscious just in case it has phenomenal qualities (or qualia for short). Phenomenal qualities are most often defined via examples, with states like seeing red and feeling pain serving as prototypical examples. For instance, Chalmers (2018, 7) writes that "it is widely accepted that seeing a bright red square and feeling pain are phenomenally conscious, but that one's ordinary background beliefs… are not."

Peressini (2014, 883) draws a further distinction here, arguing that phenomenally conscious mental states have both a qualitative character and a subjective character: "When I experience a clear blue sky, the experience has a q-character (qualitative) of 'bluishness' and an s-character (subjective) of 'being mine' in that intimate, first-person, subjective sense." Although we do not draw this distinction, our focus was on q-character. We argued, in effect, that lay people do not tend to recognize states like seeing red and feeling pain as being alike in each having a q-character, and hence do not employ the philosophical concept of phenomenal consciousness, which has q-character as a constitutive part.

The primary evidence for SM's claim that lay people generally lack the philosophical concept of phenomenal consciousness comes from the first study reported. In that study we gave either lay people or philosophers one of four vignettes describing either a normal human (Timmy) or a simple, non-humanoid robot (Jimmy) performing a color discrimination task. The descriptions for Timmy were written to either elicit the judgment that he saw red or that he felt pain. Jimmy was described in the same way, carrying out analogous behaviors. We found that philosophers tended to treat the two mental states similarly, with a clear majority ascribing each

state to the human Timmy and a clear majority denying each state of the robot Jimmy. In contrast, lay people tended to treat the two states dissimilarly. Like philosophers they tended to ascribe each state to Timmy, but unlike philosophers they tended to answer that Jimmy saw red while denying that Jimmy felt pain. Further, this finding is in keeping with the subsequent literature (see Weisman ms) and has been replicated a number of times, especially with regard to seeing colors (Sytsma and Machery 2012, Sytsma 2012, Cova et al. 2019, Sytsma and Özdemir ms1, ms2).

On the basis of these findings, as well as the results of two follow-up studies testing additional mental states (feeling anger and a variety of olfactory states), SM argue that "in clear contrast to philosophers, the folk do not seem to believe that there is something common to all of these mental states—namely that they are phenomenal" (319). In other words, participants did not tend to respond in the way that we would predict if they recognized that states like seeing red and feeling pain each have a q-character. If they recognize that both of these states have a q-character, then lay people should tend to deny both of the simple robot, just as philosophers do. That lay people happily attribute one state to Jimmy while denying the other, then, provides initial evidence that they are not treating these states as having something important in common.

2. The Ambiguity Objection

SM consider a number of objections to their conclusion. The most prominent of these is the ambiguity objection. This objection was raised by Huebner (2010), and has been pushed by Fiala, Arico, and Nichols (2014) and Chalmers (2018), in addition to Peressini. According to the ambiguity objection, the phrase "see red" is ambiguous between an *informational reading* and a *phenomenal reading*. The informational reading is less demanding, requiring simply that an

entity is able to pick up information about red things and respond appropriately to that information. In contrast, the phenomenal reading requires that the entity is phenomenally conscious and undergoes the relevant phenomenally conscious mental state (i.e., that the entity has a mental state with the relevant q-character—"reddishness" in this case). The critic then asserts that lay people draw this distinction between an informational and a phenomenal reading, perhaps implicitly, and that because the simple robot is not phenomenally conscious answer the question "Did Jimmy see red?" in terms of the informational reading. Finally, the critic asserts that people either do not draw a comparable distinction for "feel pain" or else, for some reason, default to a phenomenal reading for the question when asked whether Jimmy felt pain.

SM offer three responses to the ambiguity objection. First, we note that for this objection to avoid being blatantly ad hoc, a critic owes us an explanation for lay people's divergence from philosophers in employing the informational reading rather than the phenomenal reading of "see red." While this was not made explicit in our original paper, the issue is not simply to explain the divergence for seeing red, but to explain why there is *not* a similar divergence for feeling pain. After all, it seems that one *could* draw a corresponding informational/phenomenal distinction for "feel pain," distinguishing between an informational reading that indicates that the entity picks out relevant information about potential bodily damage and responds appropriately, and a phenomenal reading that indicates that the entity undergoes a relevant phenomenally conscious mental state. Thus, we are owed an explanation for why people employ an informational reading for "see red" *and* why they do not employ such a reading for "feel pain."

Second, we argued that our results do not bear out the ambiguity objection. We reasoned that if "see red" is in fact ambiguous between an informational and a phenomenal reading, then we would expect to see a bimodal distribution of responses, with some participants

disambiguating the question in one way and some in the other. But that is not what we found. In fact, 84.6% of non-philosophers gave an answer equal to or superior to the neutral point on our scale. Third, we noted that the explanations that non-philosophers gave for negative responses to the question "Did Jimmy see red?" did not indicate that they were calling on a phenomenal reading of the question. According to the ambiguity objection, we should expect participants who deny that Jimmy saw red to do so because they adopt a phenomenal reading of "see red" and don't believe that a simple robot is phenomenally consciousness. But the responses do not indicate that this is what is going on.

3. Peressini's Counter-response

Peressini presses the ambiguity objection against SM, offering an intriguing response to the claim that the objection is ad hoc. To do this, he contends that lay people draw a categorical distinction between *experiencers* and *non-experiencers*, classifying robots as non-experiencers. This is then taken to offer an explanation for why lay people adopt an informational reading of "see red" with regard to Jimmy. Further, Peressini provides empirical support for his claim in the first part of the study he reports. In the second part, he provides evidence that lay people have a concept of phenomenality, although it is said to diverge from the philosophical concept that SM were targeting. In this section, I detail Peressini's criticisms and respond to them directly. In the remainder of the paper, I further consider the specifics of his empirical results.

Peressini's basic objection with regard to SM is that the intelligibility of questions attributing experiential states to robots "may well presuppose that people understand robots as capable of experience in the first place" (863). This is supported by arguing that people intuitively distinguish between two different kinds of entities—experiencers and non-

experiences—with robots being thought of as non-experiences. Peressini argues that if this is correct, then "asking about perceived relative differences of mental states in robots is as meaningful as asking 'is it more efficient for a pig to fly or breath underwater?'" (863).

As detailed above, however, this was not the structure of the questions that SM asked. Rather, the analogy is more like asking two separate questions about pigs: "Can pigs fly?" and "Can pigs breath underwater?" And far from being meaningless, we would expect people to readily understand these questions and to give negative responses to both. Similarly, SM asked participants one or the other of two questions—whether the simple robot Jimmy sees red and whether the robot feels pain. We found little evidence that participants were confused by these questions, with results indicating a clear pattern of affirmation for the first question and a clear pattern of denial for the second.

Peressini's primary argument against SM is not simply that participants find the questions they asked about robots to be "senseless" (862), however, but the ambiguity objection. His claim is that assuming people intuitively draw a categorical distinction between experiencers and non-experiencers, and that they think of robots as non-experiencers, they will find it meaningless to attribute phenomenal states to the simple robot Jimmy, and *because of this* will interpret the question about seeing red in an informational sense. This argument is spelled out in more detail in Section 6.2 of Peressini's article. There, he argues that "if there is at work in the folk something like a categorical understanding of [robots] as non-experiencers… this gives us independent reason to think that there is a systematically different sense of 'seeing red' at work" (882). This offers a reply to the first response to the ambiguity objection that SM give—the response that the objection is ad hoc. Against this, Peressini argues that the distinction between experiencers and non-experiencers offers "an explanation of the ambiguity, namely, that the

informational reading of 'seeing red' is engaged in people when the entity under consideration is understood as *not being an experiencer*, and so S&M's response (1) fails" (883).

As noted above, however, SM's response is not best read as simply asking for an explanation of lay ascriptions of seeing red to the robot, but for an explanation of the overall pattern of results (although we could have been more explicit on this score). This includes explaining why people ascribe seeing red to Jimmy *and* explaining why they do not likewise ascribe feeling pain. Peressini offers an explanation of the first result, but not of the second. SM's argument is that if lay people have something like the concept of phenomenal consciousness, then they should treat these two prototypical examples of (supposed) phenomenal states similarly, but they don't. Peressini treats both types of questions about robots as being senseless on their face, and it would seem that an informational reading is available for each. So why don't people tend to answer that the robot feels pain as well as that the robot sees red?

Peressini might assert that unlike for "see red," there is no informational reading of "feel pain," as SM suggest on behalf of the critic. This assertion calls out for support, however, if the objection is to avoid being ad hoc. On the face of it, it is unclear why we should not expect a similar distinction to be drawn for "feel pain." Alternatively, Peressini might argue that people do distinguish between an informational and a phenomenal reading of "feel pain," but that the pain question doesn't trigger the same pragmatic concerns as the color question and, thus, that lay people don't shift to an informational reading. This too would call out for support, however, if the objection is to avoid being ad hoc.

Further, the strength of Peressini's explanation of attributions of seeing red to Jimmy puts corresponding pressure on the need for an explanation of participants denying that Jimmy feels pain. Peressini argues that lay people have a deep intuition that robots are non-experiencers and

that this rules out attributing experiential states to robots. And this intuition is supposed to be so strong that it overwhelmingly leads people to adopt an informational reading of "see red," since they find the phenomenal reading to be senseless when applied to a robot. But if this intuition is so deep, and if the resulting pragmatic effect is so overwhelming, we should expect similar influences to be seen for attributions of feeling pain to the robot. That is not what we find, however; instead people deny that Jimmy feels pain to roughly the same degree that they affirm that Jimmy sees red.

What's more, the above worry is exacerbated by Peressini's criticism of SM's second response to the ambiguity objection. Recall that SM's second response is that if the ambiguity objection is correct, we should expect to see a bimodal distribution of responses to the question about Jimmy seeing red, but we don't. Peressini initially reads our response as suggesting a flat distribution, but in Footnote 9 he discusses our intended point. There he argues that "the bimodal prediction would still be trumped by the systemic starting assumption that the robot is not an experiencer that triggers the informational sense of 'seeing.'" The argument appears to be simply that the effect of a robot not being an experiencer on the reading of "see red" is so strong that we don't find a notable number of participants adopting the alternative reading. This would perhaps be plausible with regard to just responses to Jimmy seeing red. However, it again rests on asserting an extremely strong pragmatic effect, which feeds into my previous point.

Peressini offers no counter to the third response given by SM. To repeat, if the results are to be explained in terms of lay people recognizing two readings of "see red," and a majority adopt the informational reading due to pragmatic considerations, then we would expect the minority who denied that Jimmy saw red to give explanations of their responses that reflect a phenomenal reading. But, again, that is not what we found. The explanations of these

participants did not suggest a phenomenal reading, but arguably "an anthropocentric reading that restricts the term to humans and animals" (311). I suppose that Peressini might respond that these explanations should be understood to be clumsily gesturing at the distinction between experiencers and non-experiencers, but the explanations do not straightforwardly suggest this distinction.

3.1 Further Evidence Against the Ambiguity Objection

I have given a number of subsequent responses to the ambiguity objection that are relevant to assessing both the ambiguity objection and Peressini's criticisms of SM. Most notably, across a series of papers (Sytsma 2009, 2010b, 2012), I have developed a fourth response to the ambiguity objection. I argue that the ambiguity objection is best seen as attaching most closely to "red" rather than "see," with the informational/phenomenal distinction corresponding with a distinction between what we might call *physical colors* (the properties of objects or light that lead to colors in us) and what we might call *phenomenal colors* (the q-qualities that we are acquainted with in having a visual experience). But there is reason to doubt that lay people typically draw such a distinction. Instead the empirical evidence suggests that they tend to treat the supposed q-qualities not as q-qualities (as qualities of mental states), but as mind-independent qualities of things in the world. Such a *naïve view* of colors does not obviously support the distinction between an informational and a phenomenal reading of "see red," however. As such, insofar as the present evidence indicates that people hold a naïve view of colors, they are not best interpreted as adopting an informational reading when they attribute seeing red to the simple robot Jimmy.

As noted above, offering an explanation of responses to Jimmy seeing red is only part of the task; we also need an explanation of why people deny that Jimmy feels pain. SM's original explanation—their *valence hypothesis*—was that people tend to hold that Jimmy is incapable of finding a state to be either pleasurable or unpleasurable and that they typically associate feeling pain, but not seeing red, with such a valence. A subsequent follow-up study (Sytsma 2012) employing larger sample sizes and a more systematic approach failed to find support for the valence hypothesis, however. In its place I offered an alternative explanation in line with the evidence just discussed, arguing that lay people also tend to hold a naïve view of pains, treating the supposed q-qualities associated with feeling pain as mind-independent qualities of injured body parts, not as qualities of mental states. This thesis was suggested by the experimental studies in Sytsma (2010) and the corpus analysis in Reuter (2011) and was backed up by further studies in Sytsma (2012). And since then the hypothesis has found a good deal of empirical support (Reuter, Phillips, and Sytsma, 2014; Kim et al., 2016; Sytsma and Reuter, 2017; Reuter and Sytsma, forthcoming; Reuter, Sienhold, and Sytsma, forthcoming).

More recently, Sytsma (2014a) responds to a version of the ambiguity objection put forward by Fiala, Arico, and Nichols (2014). Fiala et al. argue that participants in SM's study ascribed seeing red to Jimmy because they had no other way of indicating that Jimmy performed the relevant information-processing behavior.[3] They present the results of a new study suggesting that by and large lay people are unwilling to attribute seeing green to Jimmy when presented with other options, such as "Jimmy detected green." I respond by noting a number of potential issues with Fiala et al.'s study, including the puzzling finding that only a slight majority of participants (57.1%) answered that the human Timmy saw green. After systematically dealing

[3] Again, it is unclear why this would selectively apply to ascriptions of the robot seeing color, but not to ascriptions of feeling pain.

with the issues noted, I found in my final study that attributions of seeing green to Timmy rose notably (77.8%), and that they also rose significantly for Jimmy, with a clear majority now ascribing seeing green to the robot (68.8%).

Further, in Study 2 in Sytsma and Özdemir (ms2) we found a similar result using the same 7-point scale as SM, with participants still tending to attribute seeing blue to the robot (M=5.04) even when they were also asked the question "Did Jimmy detect blue?" (M=5.61). That having a way to indicate that Jimmy performed the relevant information-processing behavior does not generally lead participants to deny that Jimmy saw a color suggests against the ambiguity objection: if the proposed pragmatic effect were driving responses, we would expect the detect question to counter this, leading participants to interpret "see blue" as asking a distinct question—i.e., we would expect them to apply a phenomenal reading to the question—but they continue to attribute seeing blue to the robot.

Finally, Özdemir and I provide further evidence suggesting against the ambiguity objection (Sytsma and Özdemir ms1, ms2). In the first of these papers, we note a further manipulation suggested by Tony Jack and reported by SM:

> A critic could argue that our hypothesis predicts that if we ask ordinary people whether Jimmy *experiences red*, they would answer affirmatively. If they would answer negatively (as our critic suggests they would), then we would have to conclude that the folk conceive of subjective experience as philosophers do. (fn11)

While we echo SM's doubt that "experience" is actually used in this way by lay people (see also Sytsma 2010b), Özdemir and I argue that if lay people do in fact have a concept of phenomenal consciousness, then "experience" seems like the most likely term in English to express it. As such, we argue that if we were to replicate SM's results using this terminology, it would be compelling evidence for SM's central thesis. Focusing on seeing red, we found that participants

tended to affirm that Jimmy experienced red (M=4.73) and that responses for Jimmy weren't significantly different than responses for Timmy (M=4.98).

In Sytsma and Özdemir (ms2) we reason that if participants are reading "see red" in a purely informational sense when applied to Jimmy, then they should also tend to affirm that Jimmy sees red if the robot were to get the relevant information through another sense modality. To test this, in our first study we modified SM's original probe to specify that Jimmy's video camera isn't working, but that a smart phone is strapped to the robot that plays sounds to convey visual information about the environment, allowing Jimmy to carry out the color detection task. An analogous manipulation was carried out for Timmy. Against what we would expect if participants were generally employing a purely informational reading of "see red," we found that participants tended to deny that Jimmy saw red when he got the information auditorily (M=3.28), and similarly for Timmy (M=2.84). Comparable results were found in our second study.

The line of response to the ambiguity objection based on lay people holding a naïve view of colors and pains, as well as the new evidence from Sytsma (2014) and Sytsma and Özdemir (ms1, ms2), cast further doubt on the ambiguity objection. Not only does this evidence suggest against the ambiguity objection, but the naïve account offers an explanation of SM's findings for the full pattern of attributions to the simple robot, not just attributions of seeing red as Peressini's account does. The upshot is that Peressini's objection does not stand up well to the current body of evidence and the naïve account is to be preferred based on the present data.

3.2 A Further Objection

On the surface, Peressini seems to agree in part with SM's negative conclusion. With SM he holds that philosophers and lay people conceive of subjective experience differently, but he also

holds that lay people have a concept in the vicinity of phenomenality, although this concept is distinct from the philosophical concept. Thus, Peressini writes that "the folk *do not* think of conscious experience in terms of the standard philosophical conception of qualia/phenomenality, contrary to the assumptions of most philosophers and in keeping with Sytsma and Machery (2010)" (863). Based on statements like this, and in line with the discussion in Peressini (2017), in previous work I presented Peressini as concluding that we were "correct to deny that lay people tend to have the philosophical concept of phenomenal consciousness" (Sytsma 2016, 272). I now believe that I was mistaken.

To see why this reading was mistaken, we need to draw a further distinction within the philosophical concept of phenomenal consciousness. Above, I articulated the standard understanding of "phenomenal consciousness" in terms of an entity having mental states with q-characters. This articulation was meant to be neutral with regard to the question of whether q-characters are physical or whether they can be disconnected from the physical. In this way, the articulation I've used might be said to be *metaphysically light*: one could hold that phenomenal consciousness exists and yet hold that this is compatible with physicalism being true, that philosophical zombies are not possible, that phenomenal consciousness doesn't pose a distinctively hard problem, and so on. This corresponds with what has sometimes been called the *liberal sense* of "qualia" (e.g., Carruthers 2000, 15; see Sytsma 2010b, Section 2.2, for an extended discussion of this understanding of phenomenal consciousness).

In writing that "the folk *do not* think of conscious experience in terms of the standard philosophical conception of qualia/phenomenality," Peressini is equating the standard philosophical conception not with the metaphysically light understanding we employed, but with a *metaphysically heavy* one. Thus, he writes that "unlike philosophical qualia, the q-character

invoked by the folk in the experiment is not 'heavily' metaphysical in the sense of being thought of as something that could possibly be absent from the physical system of which it is a part" (884). As he later articulates the point:

> Call this traditional sense "m-qualia" (metaphysical), as it is the kind that can be conceived of as separable from the physical system composing an experiencer. The folk appear to employ a distinct conception of the qualitative character of experience. Call it "p-qualia" (for "physical" or "pholk"), and it has some of the properties of m-qualia: it is qualitative in that it concerns how the experience "actually appears/feels" to the subject, it is private/perspectival, and it is practically ineffable. But p-qualia are still ultimately tied to the physical. Hence, p-qualia are not the sort that "hard problem" arguments need to get off the ground. (887)

The liberal sense of "qualia" we were operating with, however, is neutral with regard to this distinction. As such, in suggesting that lay people lack the philosophical concept of phenomenal consciousness on the standard understanding—i.e., the understanding corresponding with the liberal sense of "qualia"—we were arguing that lay people by and large do not employ either the concept of m-qualia or the concept of p-qualia.

The upshot is that the conclusions drawn by SM and by Peressini are more directly in conflict than the statement that "the folk *do not* think of conscious experience in terms of the standard philosophical conception of qualia/phenomenality" would suggest. We are in agreement that lay people do not employ the concept of phenomenal consciousness underpinning there being a hard problem, but we disagree with regard to whether lay people tend to employ a more liberal sense of phenomenal consciousness. In the remainder of this paper I consider the empirical evidence that Peressini provides for this conclusion, arguing that his results are not best interpreted in terms of lay people generally employing a conception of qualia that corresponds with the liberal sense of phenomenal consciousness.

4. Peressini's Study: Experiencers

The first part of Peressini's empirical study utilizes what he describes as a direct approach to exploring the notion of "being an experiencer." It is direct in the sense of "employing from the start the notion of an experiencer," essentially "front loading" the concept (880-881). This is contrasted with the indirect approach used by SM. Peressini argues that "utilizing a more direct approach has its own distinct advantages" (867), which is true, but it also has distinct disadvantages as I will discuss below.[4]

Peressini's study was conducted on 73 university students in informal logic classes. The first part of the study opened with a description of the distinction between experiencers and non-experiencers:

> As we all know, each of us as conscious human beings have an "inner life." We are aware of things going on around us and inside our minds. In other words, there *is* something it is like to be each of us at any given moment: the sum total of what we are sensing, thinking, feeling, etc. We are experiencers.
>
> On the other hand, things like thermostats, burglar alarms, and bread machines do not have an inner life: there *is not* anything it is like to be those objects, despite the fact that they can monitor conditions around them and make appropriate things happen at appropriate times. They are not experiencers.

Participants were then asked to rate twenty entities on a 7-point scale anchored at 1 with "clearly not an experiencer," at 4 with "not sure," and at 7 with "clearly an experiencer." Entities included a normal human ("your best friend") as well as several fictional robots (Data, C3PO, R2D2) and several computer systems. There is much to say about Peressini's results, but the key point for present purposes is that while the normal human was treated as an experiencer (M=6.97), the robots and computer systems were generally treated as non-experiences, with the

[4] I've argued that in general experimental philosophers should be aiming to use a variety of methods to produce a consilience of evidence for their hypotheses (e.g., Sytsma and Livengood 2015, Sytsma et al. forthcoming). And that holds for direct versus indirect approaches. The point is that there is seldom a silver bullet when it comes to empirical work, with each method having both strengths and weaknesses.

highest rating for these entities being C3P0 (M=3.21). Coupled with the results of the second part of his study, Peressini concludes that "it is hard not to see this as evidence for a deep-seated intuition that artificial entities are categorically not experiencers" (882).

There are reasons to doubt this interpretation of the data from the first part of his study, however, both with regard to whether this distinction is categorical and whether the distinction corresponds with the sense of "experiencer" that Peressini has in mind. With regard to the claim that this distinction is categorical, it is worth noting that there is a good bit of variation in the ratings for artificial entities, although the mean for each is below the neutral point. Focusing on computers and robots, these range from M=1.48 for a computer to M=3.21 for C3P0. This might be interpreted in terms of people taking being an experiencer to fall along a spectrum—that you can be more or less of an experiencer. Further, these results might depend on the specific entities given and how they are described. More importantly, we can question whether this study taps an intuitive sense of "experiencer," and if so whether it corresponds with the sense that Peressini has in mind. In fact, there is some tension between these two claims: the more strongly people have an intuitive sense of "experiencer," the less the data warrants the conclusion that it corresponds with Peressini's articulation, and the more closely the categorizations reflect Peressini's articulation, the less the data warrants the conclusion that people are calling on a deep, intuitive sense of "experiencer" as opposed to being taught a new concept.

Peressini defends his direct approach of "front loading" the concept of an experiencer by arguing that "the central concepts involved are indeed highly recognizable by the folk" (881). But if this is correct, then there should be no need to front load the concept. In reverse, having front loaded the concept compromises our ability to conclude that this is indeed an intuitive lay concept. Here we can note that Peressini's prompt includes a number of different elements: in

introducing the category of "experiencer" he asserts that we all know that each of us is conscious and has an "inner life," that we are aware of things going on around us, that we are aware of things going on inside our minds, that there is something it is like to be each of us, and that this includes at least what we are sensing, thinking, and feeling. It could be that only some, or just one, of these elements is driving participants' responses, but they do not all clearly drive an objection against SM.

Further, Peressini provided exemplars for both categories in his prompt, giving humans as the example of experiencers and electronic artifacts (thermostats, burglar alarms, bread machines) as the example of non-experiencers. Because of this it is fair to worry that participants are simply reading the task as one of answering whether a given entity is more similar to a normal human or more similar to an electronic artifact. And, in fact, Peressini's results seem readily interpretable in this way: from lowest mean to highest mean, the ordering was Statue of Liberty, dead person, computer, computerized surveillance system, computer simulation of brain, computer simulation of brain in robot, virus, R2D2, seaweed, Data, C3PO, person in coma, Spock, human embryo, Bambi, person under general anesthesia, God, dolphin, Helen Keller, and your best friend.

That people distinguish between robots and humans is not in doubt. This is suggested not only by SM's data and my subsequent work, but by a good bit of other empirical work in experimental philosophy and psychology (see Weisman ms for a thorough survey). Further, the existing empirical work suggests that this is not best understood as reflecting a single underlying categorical distinction such as experiencer versus non-experiencer. Rather, there appear to be multiple dimensions of mind perception, with people tending to associate each with humans but only one with the robots tested. Most importantly, recent work by Weisman, Dweck, and

Markman (2017) suggests that there are three fundamental components to how adult Americans conceptualize mental life—what they term *body*, *heart*, and *mind*. The body component corresponds "primarily to physiological sensations related to biological needs, as well as the kinds of self-initiated behavior needed to pursue these needs" (11375). This component includes the capacity to experience pain. The heart component corresponds "primarily to basic and social emotions, as well as the kinds of social-cognitive and self-regulatory abilities required of a social partner and moral agent" (11375). Finally, the mind component corresponds "primarily to perceptual-cognitive abilities to detect and use information about the environment" (11375). This component includes seeing things.

Across four large-scale studies of US adults, Weisman et al. asked participants to judge whether an entity was capable of each of 40 mental capacities. Each study included a robot as one of the entities. Weisman et al. found that participants "endorsed some kinds of experience quite strongly for robots (e.g., sensing temperatures, seeing things) and rejected quite a few agentic abilities (e.g., having free will, telling right from wrong)" (11377). More generally, as indicated by their Figure 1, while people generally denied that the robot had the capacities associated with the body and heart component, they generally gave much more affirmative responses for the mind capacities. Weisman et al. suggest that "thinking about robots and other artificial intelligences as having perceptual-cognitive abilities might feel perfectly comfortable to many people, while entertaining the possibility of technological beings having social-emotional or bodily capacities may be more unsettling" (11378). This suggests against Peressini's claim that lay people employ a categorical distinction between experiencers and non-experiencers that would explain away affirmations that Jimmy saw red, which both pushes against his version of

the ambiguity objection and provides reason to suspect that his participants were not employing the distinction intended by his prompt.

The point is that while Peressini's data is clear in showing that people treat humans and robots differently with regard to the question he posed, it is not clear what drives this difference. With regard to SM, what matters is the extent to which the data suggests that lay people tend to have a concept of an experiencer tied to a conception of q-character, such that classifying robots as non-experiencers would lead them to deny that a phenomenal reading of "see red" could apply to a robot. Peressini holds that the concept of experiencer his study is tapping corresponds with Nagel's (1974) notion of "something it is like," but also holds that "it is not obvious that the problem of what makes an organism a subjective experiencer in Nagel's sense and the problem of what makes a state a qualitative state are so simply related" (864). Following on this, it is simply not clear that the results of Peressini's first study are problematic for SM's project.

The second part of Peressini's study is intended to get at whether lay people "make use of a concept of qualia in attributing 'being an experiencer'" (868), and I return to this below. Importantly, however, Peressini interprets the results as indicating that "this intuition doesn't seem to be based on the absence/presence of a (sensory) qualitative aspect, since the non-human 'intelligent creature' [in Version A of Question 7 in the second part of the study] lacks the qualitative aspect… and is still ranked as an experiencer" (882). If this is correct, then it is tough to see how the results of his first study directly support the ambiguity objection against SM. If the concept of experiencer that lay people are operating with is *not* tied to q-character, then it is not clear why people classifying robots as non-experiencers would explain their reading "see red" in an informational sense as opposed to a phenomenal sense that centrally features q-characters, as the objection contends.

Peressini could instead level a different objection, arguing that lay people (perhaps implicitly) distinguish between a *non-subjective* and a *subjective* reading of "see red," where the subjective reading is articulated in terms s-character instead of q-character (i.e., the experience belonging to the entity in "that intimate, first-person, subjective sense" instead of having a qualitative character of "reddishness"). And there is some suggestion that this is what Peressini has in mind, for example when he writes that "we can understand… the question of what makes an entity an experiencer in Nagel's sense, as fundamentally about the s-character of conscious experience, and questions about the nature of qualitative (or "phenomenological") states as fundamentally about the q-character" (883-884). Such a *subjective ambiguity objection* doesn't threaten SM's conclusion, however. The issue is that we were responding to a claim that it is pretheoretically obvious that we are phenomenally conscious in the sense of having states with q-characters, such that people should recognize that there is something important in common between states like seeing red and feeling pain (that they both have q-characters). And our results would seem to support our denial of this claim even if people distinguish between subjective and non-subjective readings of "see red."

Now, Peressini could call on the results of the second part of his study to argue that lay people *also* typically have a conception of q-character, taking this to directly contradict our denial, and then arguing that given this our results are best explained in terms of the original ambiguity objection. But this critique would not hinge on lay people distinguishing experiencers from non-experiencers in the way that Peressini lays out. It would not be the case that "if there is at work in the folk something like a categorical understanding of such entities as non-experiencers, as there seems to be, this gives us independent reason to think that there is a

systematically different sense of 'seeing red' at work" (882) that would undermine SM's conclusion.

5. Peressini's Study: Phenomenality

As I just noted, the best argument against SM based on Peressini's data comes not from the first part of his study and the distinction between experiencers and non-experiencers, but from the second part of his study where he attempts to directly assess whether lay people generally have a relevant concept of qualia (i.e., one that includes q-character). He concludes from these results that lay people do in fact typically have such a concept, although it is metaphysically light and does not support the conclusion that phenomenal consciousness poses a hard problem. Nonetheless, as detailed in Section 3, if Peressini is correct, this is still problematic for SM's conclusion, since we were targeting a liberal sense of phenomenal consciousness that covers a metaphysically light conception.

The second part of Peressini's study was conducted immediately after the first using the same 73 participants. Participants were asked 10 questions, with the first nine using a 7-point scale anchored at 1 with "certainly not," at 4 with "not sure," and at 7 with "certainly so." Again, there is much that could be said about Peressini's results. His contention that these results make it clear that "participants have a concept of some sense of qualia or 'phenomenality'" (883) rests on the results of Questions 3 and 6, however, so I will focus on these questions here.

Question 3 asks participants "Can we ever be sure that you see red the way another person does?" The mean response for this question was 3.07, indicating that participants generally gave negative answers. Peressini interprets these responses as showing that "participants' understanding of 'seeing red' involves a component that cannot be verified from a

third-person perspective" (883). He then argues that "such a private and ineffable component is typically a defining feature of qualia or 'phenomenality'" and that "there must be 'something it is like' to see red for it to be unverifiable" (883). In other words, Peressini holds that negative responses to Question 3 indicate that seeing red involves a private and ineffable component, which is tantamount to employing a relevant concept of qualia.

In line with Peressini's later discussion of "p-qualia" (the concept of qualia that he attributes to lay people), however, the sense of "private and ineffable" at issue is best understood as being qualified: p-qualia are "private/perspectival" and "practically ineffable" (887).[5] But there is a tension here. The move from negative responses to Question 3 to people recognizing that seeing red involves a private and ineffable component is more plausible for a *weak* reading of "private and ineffable," but the move from recognizing that seeing red involves a private and ineffable component to people recognizing that seeing red involves q-characters is more plausible for a *strong* reading of "private and ineffable."

Starting with the first part of the tension, if "private and ineffable" is read in a strong sense such that it coincides with in principle unverifiability, then it is a leap to interpret negative responses to Question 3 as indicating that people hold that seeing red involves a private and ineffable component. This is a leap because Question 3 can be naturally read as simply asking whether the participant is generally in a position to be sure that another person sees red the way that she does. Since we typically have to rely on relatively low information behavioral cues with regard to how people see colors, we're not generally in a position to be certain of this. A

[5] Such qualification is needed because otherwise this would seem to suggest that lay people hold a metaphysically *heavy* concept of qualia, not a metaphysically *light* concept as Peressini contends. While it is true that philosophers often hold that qualia are private and ineffable, this is not universally the case among qualia-believers. On a metaphysically light understanding of "qualia," qualia need not be either private or ineffable. That is, one could hold that qualia exist and that they can be explained in physical terms, such that they could be verified from a third-person perspective. Of course, a metaphysically heavy understanding of "qualia" would rule this out.

participant registering this fact, however, would not necessarily imply that they hold the stronger claim that how a person sees red cannot be verified from a third-person perspective in principle. By giving a negative answer to Question 3, a participant might simply be noting the in-practice difficulty of comparing abilities with regard to discerning different shades of color. To get at the stronger claim, we would need to pose a stronger question, such as asking participants whether it was theoretically possible that with enough study scientists in the future might be able to verify that two people see red in the same way.

Turning to the second part of the tension, if "private and ineffable" is read in a weak sense that falls short of in principle unverifiability, then it is a leap to infer from people indicating that seeing red involves a private and ineffable component to their recognizing that seeing red involves q-characters. The reason is that we're not generally in a position to be certain about many aspects of another person's mental states, including both those that philosophers typically treat as being phenomenal and those that they typically treat as being non-phenomenal. If negative responses to Question 3 simply indicate that people are registering the in-practice difficulty of being certain about how another person sees red, then the same type of denial would seem plausible for states that philosophers do not typically treat as being phenomenally conscious, such as thoughts or unconscious desires. While there is a good deal of debate about exactly where the division between phenomenal states and non-phenomenal states lies, it remains possible to hold that there is a component to phenomenal states that cannot be practically verified from a third-person perspective *and* that the same holds for many non-phenomenal states. In other words, a control is needed here: to conclude that people recognize something like q-character for seeing red, we would want evidence not just that participants affirm that there is

something unverifiable about seeing red, but that they deny this of states that philosophers typically treat as being non-phenomenal.

Question 6 begins with the prompt, "Imagine that a completely color blind person got an implant that encoded colors in her visual field with numbers indicating colors, so for example, the sky on a clear day was indicated with a number 1 to indicate blue." Participants were then asked, "Would such a person be able to see blue?" The mean response for this question was 3.00, again indicating that participants generally gave negative responses. Peressini argues that the "responses suggest that 'seeing blue' involves a 'qualitative aspect' or at least a constraint on 'what it would be like' for a SIL-conscious entity in order for it to 'see blue,' as opposed to detecting blue" (883). For the reasons noted above, such a constraint is not a threat to SM if "what it would be like" is cashed-out in terms of s-character rather than q-character. If the responses indicate that people hold that seeing blue involves a qualitative aspect, however, this would be problematic for SM's conclusion. As Peressini's hedge suggests, this cannot be straightforwardly inferred from responses to the question. That participants tend to deny that a completely color-blind person sees blue can be readily explained without calling on something like the concepts of qualia laid out. It is reasonable to expect that people hold that seeing a color requires visually perceiving the color, but one could hold this without commitment to such visual episodes having a q-character. As I suggest in Sytsma (2016, fn 15) one would also expect lay people to deny that a blind person sees blue when they are told that the sky is blue. But we couldn't infer from this that lay people have a concept of qualia. And, no more can we infer from people denying that a completely color-blind person sees blue when using an implant that describes colors in terms of numbers that they have a concept of qualia.

Returning to the ambiguity objection, if we interpret responses to Question 6 as indicating "a constraint on 'what it would be like' for a SIL-conscious entity in order for it to 'see blue,' as opposed to detecting blue" then Peressini's data in fact poses a difficulty for the objection along the lines of those raised by Sytsma and Özdemir (ms2) that I discussed above. What Peressini's finding suggests is that people are willing to deny that an entity that is incapable of the relevant phenomenal state sees blue. But the ambiguity objection attempts to explain away responses that Jimmy sees red in terms of people adopting an informational reading because Jimmy is incapable of the relevant phenomenal state. Why don't they do the same for the color-blind person? Peressini might respond that this is just a constraint on attributions to experiencers and doesn't constrain attributions to non-experiencers. This response does not seem to be motivated by the data, however. Even taking the divide between experiencers and non-experiencers to be deeply intuitive, it is unclear why this would predict a massive pragmatic effect for robots but not for a color-blind human. Again, the ambiguity objection appears to remain ad hoc.

6. Conclusion

Peressini offers an intriguing counter-response to SM's response to the ambiguity objection, supporting his account with new empirical evidence. His account does not explain the full pattern of SM's results, however, unlike my alternative account based on lay people tending to hold a naïve view of both colors and pains, which has also received empirical support. Further, since SM's original article there have been a number of new findings that cast further doubt on the ambiguity objection. Finally, the empirical study that Peressini bases his objections on does not straightforwardly make the case against SM and there are a number of issues with his

interpretation of the results. As such, I contend that contra Peressini, the current evidence continues to support SM's negative conclusion that the existence of phenomenal consciousness is not pretheoretically obvious. That said, further empirical work is needed concerning whether there is a lay concept of experiencers, and if so, what it looks like and what light it casts on SM's conclusion. Such work is underway and will be reported in a companion paper (Sytsma, in preparation).

References

Caruthers, Peter (2000). *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge: Cambridge University Press.

Chalmers, David (1995). "Facing up to the problem of consciousness." *Journal of Consciousness Studies*, 2(3): 200–219.

Chalmers, David (2018). "The Meta-Problem of Consciousness." Journal of Consciousness Studies, 25(9-10): 6–61.

Cova, Florian et al. (forthcoming). "Estimating the Reproducibility of Experimental Philosophy." *Review of Philosophy and Psychology*.

Fiala, Brian, Adam Arico, and Shaun Nichols (2014). "You, robot." In E. Machery and E. O'Neill (eds), *Current Controversies in Experimental Philosophy*, New York: Routledge.

Huebner, Bryce (2010). "Commonsense concepts of phenomenal consciousness: Does anyone care about functional zombies?" *Phenomenology and the Cognitive Sciences*, 9: 133–155.

Kim, Hyo-eun, Nina Poth, Kevin Reuter, and Justin Sytsma (2016). "Where is your pain? A Cross-cultural Comparison of the Concept of Pain in Americans and South Koreans." *Studia Philosophica Estonica*, 9(1): 136–169.

Machery, Edouard and Justin Sytsma (2011). "Robot Pains and Corporate Feelings." *The Philosophers' Magazine*, 1st quarter 2011: 78–82.

Nagel, Thomas (1974). "What is it like to be a bat?" *Philosophical Review*, 83: 435–450.

Peressini, Anthony (2014). "Blurring two conceptions of subjective experience: Folk versus philosophical phenomenality." *Philosophical Psychology*, 27(6): 862–889.

Reuter, Kevin (2011). “Distinguishing the appearance from the reality of pain.” *Journal of Consciousness Studies*, 18(9–10): 94–109.

Reuter, Kevin, Dustin Phillips, and Justin Sytsma (2014). “Hallucinating Pain.” In J. Sytsma (ed), *Advances in Experimental Philosophy of Mind*, Bloomsbury.

Reuter, Kevin, Michael Sienhold, and Justin Sytsma (2018). “Putting Pain in its Proper Place.” *Analysis*.

Reuter, Kevin and Justin Sytsma (2018). “Unfelt Pain.” *Synthese*.

Sytsma, Justin (2009). “Phenomenological Obviousness and the New Science of Consciousness.” *Philosophy of Science*, 76(5): 958–969.

Sytsma, Justin (2010a). “Dennett’s Theory of the Folk Theory of Consciousness.” *Journal of Consciousness Studies*, 17(3-4): 107–130.

Sytsma, Justin (2010b). *Phenomenal consciousness as scientific phenomenon? A critical investigation of the new science of consciousness*. PhD Dissertation, University of Pittsburgh.

Sytsma, Justin (2012). “Revisiting the Valence Account.” *Philosophical Topics*, 40(2): 179–198.

Sytsma, Justin (2014a). “The Robots of the Dawn of Experimental Philosophy of Mind.” In E. Machery and E. O’Neill (eds.), *Current Controversies in Experimental Philosophy*, Routledge.

Sytsma, Justin (2014b). *Advances in Experimental Philosophy of Mind*. Bloomsbury.

Sytsma, Justin (2016). “Attributions of Consciousness.” In J. Sytsma and W. Buckwalter (eds), *A Companion to Experimental Philosophy*, Blackwell.

Sytsma, Justin (in preparation). “New Experiments Concerning Experiencers.”

Sytsma, Justin, Roland Bluhm, Pascale Willemsen, and Kevin Reuter (forthcoming). “Causal Attributions and Corpus Analysis.” In E. Fischer and M. Curtis (eds), *Methodological Advances in Experimental Philosophy*, Bloomsbury.

Sytsma, Justin and Wesley Buckwalter (2016). *A Companion to Experimental Philosophy*, Blackwell.

Sytsma, Justin and Jonathan Livengood (2015). *The Theory and Practice of Experimental Philosophy*, Broadview.

Sytsma, Justin and Edouard Machery (2010). “Two Conceptions of Subjective Experience.” *Philosophical Studies*, 151(2): 299–327.

Sytsma, Justin and Edouard Machery (2012). “On the Relevance of Folk Intuitions: A Reply to Talbot.” *Consciousness and Cognition*, 21: 654-660.

Sytsma, Justin and Eyuphan Özdemir (ms1). “No Problem, Part I: Evidence that Problem Intuitions Are Not Widespread”

Sytsma, Justin and Eyuphan Özdemir (ms2). “No Problem, Part II: The Concept of Phenomenal Consciousness”

Sytsma, Justin and Kevin Reuter (2017). “Experimental Philosophy of Pain.” *Journal of Indian Council of Philosophical Research*, 34(3): 611–628.

Weisman, Kara (ms). “Machines with minds: Beliefs about the mental lives of robots and implications for human social reasoning”

Weisman, Kara, Carol Dweck, and Ellen Markman (2017). “Rethinking people's conceptions of mental life.” *PNAS*: 114(43): 11374-11379.