

1-1-2014

Blurring Two Conceptions of Subjective Experience: Folk versus Philosophical Phenomenality

Anthony F. Peressini

Marquette University, anthony.peressini@marquette.edu

Blurring Two Conceptions of Subjective Experience: Folk Versus Philosophical Phenomenality

Anthony Peressini

*Philosophy Department, Marquette University
Milwaukee, WI*

Philosophers and psychologists have experimentally explored various aspects of people's understandings of subjective experience based on their responses to questions about whether robots "see red" or "feel frustrated," but the intelligibility of such questions may well presuppose that people understand robots as experiencers in the first place. Departing from the standard approach, I develop an experimental framework that distinguishes between "phenomenal consciousness" as it is applied to a subject (an experiencer) and to an (experiential) mental state and experimentally test folk understandings of both subjective experience and experiencers. My findings (1) reveal limitations in experimental approaches using "artificial experiencers" like robots, (2) indicate that the standard philosophical conception of subjective experience in terms of qualia is distinct from that of the folk, and (3) show that folk intuitions do support a conception of qualia that departs from the philosophical conception in that it is physical rather than metaphysical. These findings have implications for the "hard problem" of consciousness.

Keywords: Experimental Philosophy; Folk Conception of Subjective Experience; Hard Problem of Consciousness; Phenomenal Consciousness; Qualia; Subjective Experience

1. Introduction

Does an intelligent robot's visual experience of red differ from its experience of physical damage? One might say, "yes, of course. Otherwise how could it take different appropriate action in each case?" But many people might find the question senseless, because they do not believe that such a robot would have conscious experience at all. Philosophers and psychologists have experimentally explored various aspects of people's understandings of subjective experience based on their responses to questions about whether such robots "see red" or "feel frustrated," but again, the intelligibility of such questions may well presuppose that people understand robots as capable of experience in the first place. The issue is that perhaps people intuitively carve up the world first and foremost not in terms of different (kinds of) mental states, but in terms of different (kinds of) entities: those entities that can have any experiences (experiencers); and those entities that cannot (nonexperiencers). If this is true—and robots are seen as not experiencing anything—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?"

In this paper, I undertake to further explore how people understand subjective experience and test a standard methodological approach to subjective experience that assumes that robots and similar artifacts can be used unproblematically to test how people think about subjective experience. To this end, instead of taking experiential states as the fundamental thing to be investigated, I take experiencers themselves as the starting point. I begin from Sytsma and Machery's (2010) results suggesting that philosophers and the folk have rather different conceptions of subjective experience. I offer new empirical results that further call into question the belief common among philosophers that there is intuitive, pre-theoretical warrant for thinking of a conscious experiencer in terms of qualitative or phenomenal states (qualia). In particular, I will argue for the following three claims:

1. One of the standard ways intuitions about conscious experience have been tested, using attributions of mental states to simple

- robots, is deeply problematic because it fails to take into account the more primitive concept of "an experiencer."
2. 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).
 3. The folk do have a conception of the qualitative aspect of conscious experience that plays a role in their understanding of conscious experience, contrary to Sytsma and Machery, and this conception is distinct from the standard philosophical conception because it is grounded in the physical as opposed to the metaphysical.

In section 2, I sketch recent work on subjective experience, focusing on Sytsma and Machery (2010). In section 3, I develop my own experimental framework and present my results in two parts with discussion in sections 4 (attributions of "experiencer") and 5 (qualia and experiencers). The final section is a general discussion of the rationale, limitations, and implications of this study for accounts of subjective experience, qualia, and the "hard" problem of consciousness.

2. Conceptions of Subjective Experience

Philosophers most often theorize subjective experience in terms of qualitative or phenomenological mental states (qualia). Qualia are understood to be the "qualitative aspects" of conscious experience; if someone is seeing a stop sign, then the particular way the red appears in his/her phenomenological field is a red quale. There are qualia associated with all sense modalities and other states like emotions, moods, etc. Standard features of qualia include the following: they appear practically ineffable, non-relational, non-public, and immediately available to the subject. Consciousness and qualia are often run together. Chalmers (1995), in setting up his "hard problem" of consciousness, moves between Nagel's (1974) "something it is like" notion of an experiencer and the notion of qualia. But 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. In recognition of this, experimental work focusing on how people actually understand subjective experience has emerged in the last few years.

Such studies can be found in the work of Robbins and Jack (2006) on the conditions necessary for the folk to understand something as having subjective experience, or as they put it, taking the "phenomenological stance" toward the entity. Gray, Gray, and Wegner's (2007) empirical study of "mind perception" involved a systematic examination of the factors involved in how people recognize minds in which they documented two major dimensions to such perception, "experience" (feeling pain, hunger, anger, etc.) and "agency" (self-control, memory, morality, etc.). They had participants compare a host of entities (e.g., frog, robot, God, infant, etc.) with respect to the entities' capacity to have mental states and found that the states fell into natural groups along the dimensions of "experience" and "agency." Also, Knobe and Prinz (2008) have studied how "the folk" understand subjective experience by analyzing how people differentially attribute phenomenological and non-phenomenological mental states to group agents (e.g., corporations); they argued that the unwillingness of people to attribute phenomenological states to group entities (in contrast to non-phenomenological states) indicate that the folk recognize such a difference. But none of these studies target precisely subjective experience in the broad sense of "phenomenality" or the general qualitative nature of subjective experience. Robbins and Jack (2006), with their focus on "hedonic value," miss the most general sense of subjective experience because they do not include in their study the mundane subjective aspects of perceptual states lacking in hedonic value, e.g., the particular way the brown color of the bricks appears to me or the particular way the hum of traffic in the background sounds. Similarly, Gray et al. (2007), while using the term "experience" in their study, do not use it in the philosophical sense of general subjective experience because they do not include perceptual experience in their sense of it. Finally, Knobe and Prinz (2008) do actually target subjective experience in its most general phenomenal sense, but their choice to use in their study collective entities like corporations, which can be shown to differ functionally and behaviorally from individuals, confounds their conclusions. I refer the reader to Sytsma and Machery (2009; 2010, pp. 302–305) for further critical details. In what follows, I will make

primary use of the critical framework of Sytsma and Machery (2010; henceforth "S&M") to contextualize and motivate this study, since they do indeed address the question of subjective experience in its most general sense.¹

S&M set out explicitly to examine whether the "folk understanding" of subjective experience has been "read off" successfully by philosophers. S&M point out the pervasive philosophical assumption of a deep link between subjective experience and phenomenality. They write in the first sentence of their first section, "for most contemporary philosophers, subjective experience is characterized by its phenomenality" (2010, p. 300). S&M's approach involves finding experimental settings that cleave between the "philosophical" conception and a distinctive folk conception. S&M make use of a robot/human vignette as the backbone of their experiment, analyzing the differential patterns of the attribution of mental states by the participants to the vignette's main character, either a robot or a human. The independent variables were (1) whether the main character in the vignette was a human or robot and, (2) which mental state they were asked to attribute (or not) to the main character.

In the first of their three studies, participants were randomly assigned to one of four vignettes varying the subject between human/robot (H or R), and varying the mental state between seeing red/feeling pain (SR or FP). In the first version, the robot named Jimmy (or human named Timmy) is instructed to move a red box from a group of variously colored boxes. After it/he successfully moves the box, participants were asked whether Jimmy/Timmy "saw red." In the second version, Jimmy/Timmy would again be instructed to move the red box but would receive an electrical shock after picking up the box, immediately drop the box, and move quickly away from it. Participants were then asked whether Jimmy/Timmy "felt pain." Participants answered the question with a number on a seven-point scale with 1 $\frac{1}{4}$ clearly no, 4 $\frac{1}{4}$ not sure, and 7 $\frac{1}{4}$ clearly yes. The participants were also pre-screened to determine whether they were philosophers (had graduate training or were majoring in philosophy), and the results were further broken down by the division of participants into philosophers and non-philosophers. This constituted a test of the hypothesis that non-philosophers and philosophers both conceive of conscious experience (at least implicitly) in phenomenological terms,

the rationale being that for those thinking in terms of phenomenological consciousness, the attribution of perceptual experience (red) and bodily sensation (pain) would go together. While the prediction for philosophers was borne out, it was not for non-philosophers; there was a significant difference in that non-philosophers were willing to attribute "seeing red" to the robot and philosophers were not. Neither group attributed "feeling pain" to the robot, and they both attributed both pain and seeing red to the human. These results suggest to S&M that "ordinary folk" conceive of subjective experience differently from philosophers.

In the second study, S&M proposed to test (and ultimately reject) a putative explanation of the attribution behavior of the non-philosophers that avoided the conclusion that the difference was due to non-philosophers not recognizing the phenomenological features of experience. The alternative explanation is that there is a relevant difference between internal senses (pain, emotions, etc.) and external senses (perceptual experience), and that the non-philosophers were willing to attribute external senses (seeing red) to the robot, and not internal ones (feeling pain), but nonetheless understood sensing phenomenally. In this experiment, participants were again divided into four groups; the scenarios were modified to be about moving a box of bananas (detecting it by odor from boxes of other distinctively smelly things), and the robot/person was either successful or frustrated in moving the box. Participants were asked whether the person/robot smelled bananas in the successful scenarios and whether the person/robot felt angry in the frustrating scenarios. If the internal/external hypothesis were correct then the non-philosophers should be willing to attribute smelling bananas to the robot (like they did with "seeing red"), but not with "feeling anger." The results did not bear this out: "smelling bananas" was neither attributed (nor denied) to the robot. The participants seemed ambivalent and divided as to whether to attribute "smelling" to the robot. Thus the results were interpreted as refuting the claim that the internal/external sense hypothesis governed attributions. One possibility S&M acknowledge is that sense modalities could vary in how external they are (e.g., vision being very external and smelling less external), and thus a refined internal/external hypothesis, one that recognizes sense modalities as varying between more internal and less internal, might still be consistent with the results.

In the final study, S&M introduce a factor that they suggest does explain the preceding attribution behavior, namely that mental states with an "affective valence" (states that one either wants to be in or wants to avoid being in) are ones that are associated with subjective experience. As S&M put it, "we hypothesize that it is not whether a mental state is the product of the external senses that matters for the folk understanding of subjective experience, but whether they associate that state with some hedonic value for the subject" (2010, pp. 314–315). To test this positive account (and rule out the refined internal/external hypothesis), they ran a similar scenario test with a robot/human and three different olfactory cues: familiar pleasant (banana); familiar unpleasant (vomit); and unfamiliar and presumably valence neutral (isoamyl acetate). Again, the results indicated that the participants were willing to ascribe all smells to the human, but only the unfamiliar cue (isoamyl acetate) to the robot. The scores for attributing familiar smells to the robot were not significantly different from "not sure," so while participants failed to attribute familiar smells to the robot, neither did they deny that the robot smelled the familiar objects. S&M argue that these results favor their valence proposal over the alternative interior/exterior distance proposal, because their valence account predicts the observed response pattern, which falls along different valence lines. The alternative hypothesis cannot explain the difference because the smells were all from the same sense modality and hence did not involve any difference in the "interior distance."

S&M view the results of their studies as constituting preliminary evidence that phenomenality does not figure into how non-philosophers attribute states like seeing color, hearing a sound, feeling a pain, etc. In their final section, S&M argue that since non-philosophers do not recognize these diverse mental states as united by having phenomenological properties, it follows that phenomenological experience cannot be taken to be the manifestly obvious feature of our mental life that the philosophers pushing the "hard problem" of consciousness need it to be (2010, p. 321). S&M are advancing an important skeptical line against the "hard problem." In particular, they have raised serious doubts about whether non-philosophers have the same conception of phenomenality that philosophers do. There are, however, several problematic aspects of their study that threaten to undermine the support for their conclusion.²

The first concern is that asking participants to attribute mental states to the robot character might not be an effective way to test attribution behavior. Robots and other artifacts like computers and artificial intelligence may from the start embody, for us, a certain "valence" in that they produce in us an aversive feeling of uncanniness when presented as experiencers. Gray and Wegner (2012) have documented an "uncanniness" effect of just this kind: machines that appear to experience states like pain, hunger, fear, and other emotions strike people as uncanny. S&M suggest that the link between subjective experience and phenomenality "only becomes obvious as one is trained into a particular way of thinking about the mind" (2010, p. 323); I suspect that such training also obviates our natural sense of "uncanniness" and explains why philosophers' intuitions seem to point in different directions. I return to this below, but the upshot is that for some reason, one certainly worth exploring, the folk begin with rather deep and pervasive intuitions that artifacts (as opposed to "natural" life) categorically cannot be experiencers and that this governs their attribution behavior.

A further concern is that S&M's approach does not push the empirical questioning back quite far enough. Their results, though preliminary, do seem to provide serious reason to doubt the intuitive grounding of the received philosophical treatment of subjective experience in terms of phenomenality. But while the folk may not conceive of subjective experience in phenomenological terms, understood in the usual philosophical way, it is moving too quickly to dismiss altogether a phenomenological component in mental state attribution. Putting "phenomenality" to more direct empirical test is called for to shed light on the folk conceptual dynamics around subjective and qualitative experience.

Finally, note that S&M use what might be termed an "indirect" approach to test how the folk think of subjective experience: they test the willingness of participants to attribute various mental states, which we assume include phenomenality or other subjective features of interest, to a robot and then "infer" to the best explanation of the attribution behavior in terms of the participants' conceptual scheme regarding subjective experience. This approach is indirect in that it requires inference back from attribution tendencies to how the participants are conceiving of things in order to explain the attribution behavior. It has the virtue of not depending on the participant's ability

to articulate implicit understandings that govern their attribution behavior, and of course, is less susceptible to the bias that may occur with more direct approaches. I will argue below, however, that utilizing a more direct approach has its own distinct advantages.

3. Experimental Design

Departing from S&M, the theoretical framework in this study centers on the notion of "being an experiencer" at a creature level rather than on experiential mental states. The explanandum is subjective conscious experience, understood in Nagel's "something it is like" (SIL) sense. An entity is an experiencer when there is "something it is like" to be that entity. I depart from S&M in that I do not collapse being an experiencer (having SIL) into having a particular mental state (or set of states) with the second-order property of "there being something it is like to have it." For purposes here, the reigning philosophical account of qualia as the phenomenological properties of mental states is taken as a starting point, and the central question is the extent to which folk conceptions agree with this account. (I move back and forth between the terms "qualia" and "phenomenality," depending on context, with the understanding that qualia are the phenomenological properties of mental states.)

The research questions for this study are (1) the ways in which the application of the notion of "being an experiencer" in Nagel's SIL sense depend on psychological, biological, behavioral, or computational factors of the putative experiencer, and (2) whether/how subjects make use of a concept of qualia in attributing "being an experiencer." The study itself involved a series of questions in two parts, the first exploring how subjects attribute "being an experiencer" and the second probing their understanding of qualia and the extent to which it affects their attributing "being an experiencer." In part 1, participants were asked to assign a degree of confidence to whether various entities are experiencers, and in part 2, participants were asked degree of confidence questions concerning being an experiencer in situations in which qualia and other related factors were manipulated. Instructions were designed to attune subjects to the anchoring cases for "being an experiencer," namely the sense of being an experiencer we have from our own case on one end, and on the

other end, simple artifacts generally understood to be lacking inner experience.

4. Experimental Study Part 1: Attributing “Being an Experienter”

This first part of the study tested the attribution behavior of subjects with respect to particular kinds of entities. It was designed to test both whether the two anchoring cases in the instructions were comprehensible to participants, and also what factors play a role in how “being an experienter” is employed by participants.

4.1. Participants

Participants (N = 73) were university students drawn from three sections of an informal logic class taught by the same instructor. The philosophical content of the course is minimal. The gender breakdown was 47% (34) male, 53% (39) female.

4.2. Materials and Procedures

Participants completed a questionnaire at the beginning of a class. The instructor administered the questionnaire and was not informed of the research questions. The participants were given the following prompt:

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 these objects, despite the fact that they can monitor conditions around them and make appropriate things happen at appropriate times. They are not experiencers.

They were then presented with a list of twenty items and asked to indicate, for each item, whether that item was an experienter using

a seven-point Likert scale anchored with 1 $\frac{1}{4}$ clearly not an experiencer, 7 $\frac{1}{4}$ clearly an experiencer, and 4 $\frac{1}{4}$ unsure. The twenty items they evaluated were:

1. a person in coma _____.
2. a virus _____.
3. Spock (from Star Trek) _____.
4. seaweed _____.
5. Helen Keller (when alive) _____.
6. a complicated computerized surveillance system visually monitoring a house and also monitoring the sounds, temperature, and odors in order to detect intruders or environmental changes and correct them _____.
7. C3PO (from Star Wars) _____.
8. a dead person _____.
9. the Statue of Liberty _____.
10. a computer program exactly simulating the behavior of your neurons _____.
11. a dolphin _____.
12. your best friend _____.
13. a computer _____.
14. Data (from Star Trek) _____.
15. Bambi _____.
16. a human embryo _____.
17. God _____.
18. a computer program exactly simulating the behavior of your neurons in a robot _____.
19. R2D2 (from Star Wars) _____.
20. a person under general anesthesia _____.

Participants were also instructed to explain their responses if necessary.

4.3. Results and Discussion

The mean values and standard deviations for the twenty items are reported in table 1. Results are graphed with a 99% confidence interval in figure 1. Table 1 contains the results of a one-sample t-test for each of the items testing the null hypothesis that the mean equals 4 (the "unsure" response). The t-tests provide rough indication of which entities were considered experiencers. In all the tests but one the results were significant at $p, 0.001$; the test for the entity Spock

was significant at $p, 0.01$. As a result, 9 of the 20 entities were ranked as experiencers and 11 were ranked as nonexperiencers; there were no entities with a mean response of "unsure."

In descending order of means, the experiencers were: your best friend; Helen Keller; dolphin; God; person under general anesthesia; Bambi; human embryo; Spock; and person in coma. The entities identified as non-experiencers (again in descending order of means) were: C3PO; Data; seaweed; R2D2; a virus; a computer simulation of your brain in a robot; a computer simulation of your brain; a computerized house surveillance system; a computer; a dead person; and the Statue of Liberty.

The mean scores for the twenty items were tested using an ANOVA with repeated measures and a Greenhouse-Geisser correction (Mauchly's Sphericity test was significant; $p, 0.001$, $\eta^2 = 0.656(189)$, $e = 0.448$). The mean scores for being an experiencer were statistically significantly different; $F(8.51, 536) = 122$, $p, 0.001$, $h^2 = 0.659$. A Bonferroni adjustment for all 190 multiple pairwise comparisons would have involved an excessive loss of power ($\alpha = 0.05/190, 0.0003$), so instead a False Detection Rate method (FDR) was employed with $\alpha = 0.007$ used for post-hoc comparisons (see table 2).³ Three discrete breaks occurred (a) between best friend (the experiencer with the highest mean) and Helen Keller; (b) between the experiencer with the lowest mean (person in coma) and the non-experiencer with the highest mean (C3PO), naturally dividing the experiencers from the non-experiencers; and (c) between the virus and the robot with a computer brain simulation. Thus a natural partition using the FDR subsets involves subsets 1, 2 vs 3, 5, 6, 7 vs the Statue of Liberty. In this partition, the first subset consists of the anchoring case (best friend), the second subset of slightly more complicated people or people-like entities (higher animals, anesthetized and impaired people, God), the third subset of marginal experiencers (fictional animals and aliens, embryo, person in coma), the fourth subset of the high non-experiencers (fictional anthropomorphized robots, virus, seaweed), and finally in the fifth subset were the low non-experiencers (explicitly computerized items, corpse, Statue of Liberty).

A Principle Component Analysis (PCA) was employed to transform the original variables onto uncorrelated components to simplify the data structure, eliminate redundant and unreliable

descriptors, and reveal interpretable factors. The relatively low sample size ($N = 73$) in this study indicated that the number of variables (20) ought to be reduced to 12 or 13 so that the standard ratio of 5:1 for sample size to variable be maintained. Only 3 of the 20 items failed to correlate at least 0.3 with at least one other item (embryo, God, corpse) indicating factorability. Individual measures of sampling adequacy on the full 20 variables indicated four variables with low (< 0.5) measures: corpse; dolphin; God; and embryo. The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.689, below the target value of 0.7, and Bartlett's test of sphericity was significant ($\chi^2(190) = 546, p, 0.001$).

After removing the four variables with low sampling adequacy, the PCA method was used and six components converged with eigenvalues > 1 , explaining 68% of the total variance. Inspection of the component matrix indicated that one component depended on only the two variables with most extreme means and smallest variances (best friend, Statue of Liberty) and one component on only one variable (Helen Keller). After removing these three variables and extracting four components, the rotated component matrix revealed one variable (computerized house) with non-zero loadings on more than two components; this variable was removed. The resultant extraction consisting of the four components explaining 73% of the variance was rotated (Varimax with Kaiser Normalization) so that the rotated factors cumulatively explained 28%, 16%, 14%, and 14% of the total variance, respectively. In the final extraction, all individual variable adequacy scores were above 0.5, the group Kaiser-Meyer-Olkin sampling adequacy score was above 0.7, and all the communalities were above 0.5. Bartlett's test of sphericity was again significant ($\chi^2(66) = 363, p, 0.001$). The rotated component matrix revealed the near simple structure shown in table 3.

The first and primary component should be interpreted as whether the entity in question is an artifact, as it loaded on all and only the artifact variables. The second component loaded significantly ($|r| > 0.5$) on the two "incapacitated" humans variables, but interestingly, also saliently ($|r| > 0.4$) in the opposite direction (negatively) on the two variables for computerized human brain/body simulations. This result is suggestive of Gray and Wegner's (2012) uncanniness findings: humans who lack importantly human qualities and artifacts that appear to have such qualities strike people as

uncanny. Accordingly, component 2 can be interpreted as a measure of uncanniness in roughly this sense.⁴ The third component loaded significantly on nonhuman living entities (seaweed and virus), and next (though not saliently) on Bambi. Finally, the fourth component loaded significantly on Spock and Bambi and seemed interpretable as a measure of human-likeness or anthropomorphicity. The internal consistency for the four factors was tested using Cronbach's alpha. The descriptive statistics for the components, Artifact (N = 6, M = 2.35, SD = 1.26, α = 0.85), Uncanniness (N = 2, M = 5.21, SD = 1.34, α = 0.74), Living (N = 2, M = 2.79, SD = 1.52, α = 0.71), and Anthropomorphic (N = 2, M = 4.94, SD = 1.91, α = 0.74) indicated "good" internal consistency.

The results indicate that the assumptions behind the anchoring cases of being an experienter matched the participants' own understanding: the significantly highest/lowest mean answers and lowest standard deviations were for "best friend" and "Statue of Liberty," as predicted. Further indication that the participants' understanding of "being an experienter" fit well with Nagel's "something it is like" articulation is that in addition to ordinary and variously incapacitated people, animals too were readily accorded "experienter" status, while at the other end, dead people and artificial entities described as "computers" are clearly not conceived of as experiencers in this sense. Fictional aliens, fictional anthropomorphized intelligent robots, embryos, people in comas, viruses, and seaweed were less conclusively non-experiencers.

5. Experimental Study Part 2: Qualia and "Being an Experienter"

The second part of study probed how the concept of qualia is related to attributions of experienter. In particular, it attempted to investigate whether subjects employ the notion of qualia and if so, what sense of qualia subjects employ, and how the absence or attenuation of qualia affects their willingness to attribute "being an experienter."

5.1. Participants

The participants consisted of the same university students (N ¼ 73) drawn from three sections of an informal logic class taught by the same instructor; the second part of the study was conducted during the same session as the first.

5.2. Materials and Procedures

Participants completed a second questionnaire immediately following the first. Again, the instructor administered the questionnaire and was not informed of the research questions. The prompt and questions themselves from part 1 were intended to further attune participants to the notion of being an experiencer in Nagel's sense. The term 'qualia' was not defined directly nor used in any of the questions. Participants were simply given the following directions:

Please answer the following questions using the scale below.

1	2	3	4	5	6	7
certainly not			not sure			certainly so

They were then asked ten questions (several requiring two responses). All but the last one involved utilizing the above Likert scale; question 10 required a written response. Two of the questions were asked in two different ways (questions 2 and 7). The ten questions in order were:

1. Imagine a medical procedure that would remove your inner experience without affecting your brain, so from the outside you would remain unchanged physically and behaviorally. Do you think such a procedure is possible? _____.
2. [Version A] Could a robot **EVER feel** anxious? _____.
[Version B] Could a robot **EVER be** anxious? _____. [Bold emphasis indicates difference in versions.]
3. Can we ever be sure that you see red the way another person does?_____.

4. Imagine a person who feels no sense of guilt or right or wrong about any actions whatsoever.
Is such an entity possible? _____.
Would such a person be an experiencer? _____.
5. Imagine a person without any emotions whatsoever.
Is such an entity possible? _____.
Would such a person be an experiencer? _____.
6. 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.
Would such a person be able to see blue? _____.
7. [Version A] Imagine **another kind of intelligent creature (from a society much like ours)** whose experience didn't consist of the subjective feels, tastes, colors, sounds, etc., but rather of objective measurements of pressure, direction, chemical composition, light frequency, physically helpful/harmful environmental factors, etc. [Bold emphasis indicates difference in versions.]
[Version B] Imagine **an intelligent robot** whose experience didn't consist of the subjective feels, tastes, colors, sounds, etc., but rather of objective measurements of pressure, direction, chemical composition, light frequency, physically helpful/harmful environmental factors, etc. [Bold emphasis indicates difference in versions.]
Is such an entity possible? _____.
Would such an entity be an experiencer? _____.
8. Imagine a person physically and behaviorally identical to you in all ways but who had no inner experience at all.
Is such a person possible? _____.
Would such an entity be an experiencer? _____.
9. Imagine a person whose senses were destroyed by disease, but who had been given artificial senses that reported directly to their speech center things like "wall up ahead" or "bird singing ten feet to right."
Would such an entity be an experiencer? _____.
10. Imagine that Dan, a professional coffee taster for years, has recently begun to fail competency tests that he used to pass all the time. He has begun to confuse coffee type X with coffee type Y, and vice versa. For him coffee type X now tastes just like coffee type Y used to and vice versa. List all the explanations you can think of that might explain this. (Continue on back if necessary.)

Participants were randomly assigned version A (N ¼ 36) or B (N ¼ 37) of the instrument that corresponded to the versions of questions 2 and 7 indicated above. The two versions of question 2 simply varied in the wording of whether a robot could "EVER feel anxious" (version A) or

"EVER be anxious" (version B).⁵ On question 7, the feature that varied was whether the entity was an "intelligent creature (from a society much like ours)" (version A) or whether the entity was "an intelligent robot" (version B).⁶

5.3. Results and Discussion

The mean values and standard deviations for questions 1–9 from part 2 are given in table 4, and graphical results with a 99% confidence interval are given in figure 2. Table 4 also contains the results of a one-sample t-test for each of the items testing the null hypothesis that the mean equals 4 (the "unsure" response). Again, the t-tests provide rough indication of whether the mean response indicates an affirmative, negative, or unsure answer.

The absent qualia/zombie questions 1 (M $\frac{1}{4}$ 2.30, SD $\frac{1}{4}$ 1.60) and 8a (M $\frac{1}{4}$ 2.55, SD $\frac{1}{4}$ 1.78) indicate that non-philosophers do not think such cases are possible, as the means differed significantly from the "unsure" level 4, with $t(72) \frac{1}{4}$ 29.09, p , 0.001, $d \frac{1}{4}$ 1.06 for question 1 and $t(72) \frac{1}{4}$ 26.97, p , 0.001, $d \frac{1}{4}$ 0.82 for question 8a. The question exploring intuitions about having a sense of guilt and right/wrong (question 4b, M $\frac{1}{4}$ 5.25, SD $\frac{1}{4}$ 1.85) indicated that such a sense is not taken as necessary for being an experiencer, $t(72) \frac{1}{4}$ 5.77, p , 0.001, $d \frac{1}{4}$ 0.67.

On the other hand, having emotions seems to neatly divide non-philosophers: the responses indicated "unsure" (question 5a: M $\frac{1}{4}$ 3.73, SD $\frac{1}{4}$ 2.31; 5b: M $\frac{1}{4}$ 4.00, SD $\frac{1}{4}$ 2.24) with both means failing to differ significantly from the level of 4. The frequencies of responses for whether such a person would be an experiencer further bear out this uncertainty, i.e., 44% indicating "no," 41% indicating "yes," and 15% "unsure." The numbers were similar for whether these persons are possible. Question 3, which asked whether we could be sure another person sees red the way we do (M $\frac{1}{4}$ 3.07, SD $\frac{1}{4}$ 2.00), indicated that participants do not think one could be sure that red things appear the same to everyone; $t(70) \frac{1}{4}$ 23.91, p , 0.001, $d \frac{1}{4}$ 0.46. The questions concerning unusual sensory qualia (questions 6 and 9) split in an interesting way. Question 6, concerning whether a color blind person with 1's in his/her visual field indicating blue could "see blue," indicated that non-philosophers thought "no" (M $\frac{1}{4}$ 3.00, SD $\frac{1}{4}$ 1.97 with $t(71) \frac{1}{4}$ 24.30, p , 0.001, $d \frac{1}{4}$ 0.51). Yet with respect

to whether a person whose sense modalities were wired directly to their speech centers (question 9, $M = 4.75$, $SD = 1.88$), the consensus was that the person would be an experienter, $t(72) = 3.42$, $p = 0.01$, $d = 0.40$. Thus folk attribution of sensory experience does seem to require a qualitative aspect in order to phenomenologically "see" as opposed to informationally "detect" blue.

The mean scores for the ten version-independent items (1, 3, 4a & b, 5a & b, 6, 8a & b, 9) were also tested using a one-way repeated-measures ANOVA. Mauchly's Sphericity test indicated the assumption of sphericity was violated, $\chi^2(44) = 122.2$, $p = 0.001$, so a Greenhouse-Geisser correction of $\epsilon = 0.732$ was used. The ANOVA shows that the mean scores are significantly different; $F(6.59, 454.6) = 33.4$, $p = 0.001$, $\eta^2 = 0.326$. A Bonferroni adjustment for the 45 multiple pairwise comparisons was used for post-hoc pairwise comparisons with $\alpha = 0.05 / 45 = 0.00111$. Comparisons of particular interest were among the six parallel questions asking about the possibility and experienter status of a person without a sense of guilt (4a, 4b), a person without any emotions (5a, 5b), and a person identical to the participant but without inner experience (8a, 8b); the results indicated no significant difference between the means of the possibility and experienter status of each of these pairs, and each of the means for the possibilities were mutually significantly different (4a – 5a – 8a), as were the means for the experienter status (4b – 5b – 8b). The two absent qualia/zombie possibility questions (1, 8a) showed no significant difference from each other, and both differed significantly from the means for the emotionless experienter question (5a, 5b), which again did not differ significantly from the unsure response level of 4.

The results for the version-dependent questions (2, 7p, and 7e) are also given in table 4 and the results are graphed with a 99% confidence interval in figure 3. Table 4 also contains the results of a one-sample t-test for each version (A and B) of the three items (total of six), testing the null hypothesis that the mean equals 4 (the "unsure" response). In terms of the rough direction of the responses, both versions of the "robot being anxious" question (2) had means indicating "no" (A: $M = 2.86$, $SD = 2.00$; B: $M = 2.41$, $SD = 1.82$) as t-tests were significant; A: $t(35) = 23.41$, $p = 0.01$, $d = 0.57$; B: $t(36) = 25.34$, $p = 0.001$, $d = 0.88$. In contrast, both versions of the possibility of an "objective experienter" question (7e) had means

indicating "yes" (A: $M = 4.86$, $SD = 1.61$; B: $M = 5.16$, $SD = 2.06$); t-tests were again significant; A: $t(35) = 3.22$, $p = 0.01$, $d = 0.54$; B: $t(36) = 3.43$, $p = 0.01$, $d = 0.56$. Finally, an interesting divergence was found in the means of the experienter status question (7e) of the difference versions of the "objective experienter." The means for the intelligent creature version (A: $M = 5.08$, $SD = 1.86$) and the intelligent robot version (B: $M = 2.70$, $SD = 1.70$) were significantly different from the "unsure" response (A: $t(35) = 3.450$, $p = 0.01$, $d = 0.58$; B: $t(36) = 24.65$, $p = 0.001$, $d = 0.76$), but in opposite directions: the creature was understood to be an experienter and the robot as a non-experienter.

An analysis of variance showed that, indeed, there was a significant effect due to the kind of intelligent entity (A: natural "creature" versus B: artificial "robot") for whether it was considered an experienter in question 7e; $F(1,71) = 32.7$, $p = 0.001$, $\eta^2 = 0.315$ (Levene's Test indicated equal variances; $F = 0.002$, $p = 0.96$). The ANOVA found no effect due to kind of entity on the possibility of experience (7p), $F(1,71) = 0.483$, $p = 0.49$, $\eta^2 = 0.006$, with Levene's Test non-significant at $F = 3.09$, $p = 0.08$. Finally, in keeping with S&M's results, there was not a significant difference between the wordings (being anxious or feeling anxious) in question 2, $F(1,71) = 1.04$, $p = 0.31$, $\eta^2 = 0.014$, and Levene's Test non-significant at $F = 0.69$, $p = 0.41$.

The final question, a variation of Dennett's (1988) Chase and Sandborn coffee taster "intuition pump," tested per S&M's (2010, p. 323) suggestion whether non-philosophers might be readily nudged into recognizing a perhaps latent conception of qualia. In it, participants were asked to list all possible explanations for the coffee taster's "interchanged qualia." The 123 responses offered were categorized into 15 groups with results in table 5 and figure 4.

The vast majority (111, 90.2%) of all responses fell into one of the following four non-phenomenological categories: taste bud issues (explanation types 1, 2, 3, 4); old age/disease (explanation types 7, 8); brain/nerve damage (explanation type 10); and too used to the taste (explanation type 9). There were only seven (5.7%) of the 123 explanations that offered reasonably clear phenomenological responses; these explanations fell into three types (11, 12, 13): the memory of the taste changed (two participants); the perception of the

taste changed (one participant); and altered inner experience (four participants). Notably, five of the seven responders offering a phenomenological explanation had indicated that they had at least one prior philosophy class. Finally, there were four responses (4%) of self-deception/doubt (explanation types 5, 6) that one could possibly interpret as phenomenological. Especially given that the previous questions and prompts in the experiment asked participants repeatedly to think about "inner experience" and its variability, this direct approach to awaking putative latent intuitions concerning qualia/phenomenality provided no evidence that any such latent concepts exist.

6. General Discussion

In this section, I discuss the rationale for the experimental framework (section 6.1), the implications of the results for experimental methodology and general understanding of putative "artificial experiencers" (section 6.2), the relationship between conscious subjective experience and qualia (section 6.3), limitations of the study (section 6.4), and implications for the "hard" problem of consciousness (section 6.5).

6.1. Theoretical Framework: Direct versus Indirect Design

Part of the experimental design involves assessing subjects' direct responses. For example, the participants in the study were asked to assign a degree of certainty to whether an entity under consideration is an experiencer. This direct approach, employing from the start the notion of an experiencer as it does, can be thought of as "front loading" a concept. However, no contentious philosophical assumptions were made about it; the experiment simply explores the commonly held notion that certain things are experiencers and others are not.

A natural concern when one attempts experimentally to test direct responses is bias. The problem of bias in experiments with direct questions may come up in various ways, including in particular demand characteristics, the experimental artifact introduced

when participants behave out of a perhaps implicit and unconscious interpretation of the purpose of the experiment. Another possible source arises from the difference between the employment of a concept and making judgments about such employment. In this setting, however, it is not clear that demand characteristic bias will be a factor since the questions do not concern "socially loaded" areas like sexual practice, drug use, or racial attitudes in which demand bias is clearly a concern.⁷ Similarly, while a dualprocess background hypothesis might suggest that a higher-level report is distorting the lower-level process of interest, given that this study's interest is not necessarily in such low-level processes, this sort of bias is not a clear concern either.

This is not to say there is not the possibility for bias arising out of the explicit structure of part of the experiment, but rather that the likelihood is not a great concern here because (1) low-level distortions from a dual process situation are not particularly germane since the target is a higher conceptual dynamic, and (2) demand characteristics, which typically play out in attitudinal studies with a social expectation valence, are not clearly indicated because the topics involved are not generally socially contentious. It was not feasible to implement a post-experimental questionnaire, which is a standard technique to test for demand characteristics. I did, however, employ a measure to counter such bias: the administrator of the experiment was not informed of the research hypothesis or topic.

There is a complementary concern when one attempts experimentally to test attribution behavior with a concept like "experienter," namely, that subjects may simply lack an understanding of the concepts required (subjective experience, qualia, etc.) to make responses meaningful. But in this case, concern about subjects not having the requisite concepts may be obviated by the fact that the central concepts involved are indeed highly recognizable by the folk, namely the "how colors look to each of us" sense of qualia and the "something it is like to be" sense of an experienter. And as anyone who has taught these concepts in philosophy or psychology can attest—there is enough of an understanding in the folk of the "what it is like to see red" notion to immediately grasp the epistemological conundrum to which it gives rise. What is more, the case at hand is different from cases like asking subjects whether

there can be truths in math that cannot be proven or whether the largest prime number is odd or even.⁸ Of course, in such cases one may not conclude much of anything about whether the folk agree with Gödel's Incompleteness Theorem or Euclid's proof that there are an infinite number of primes. But this study targets how folk understandings of a *relatively ordinary* concept (not precisely defined terms like even, odd, prime, proof, or truth) compare with "expert" understandings of these same ordinary phenomena, not results deduced from propositions about these phenomena. In asking participants to employ their concept of "how red looks" to novel and unusual situations, the understanding is that how they apply it will shed light on its contours as a folk concept, not whether they actually disagree with Jackson's Mary argument.

6.2. Artificial Experiencers

The results from parts 1 and 2 make a clear case for a categorical resistance in the folk to attributing experiencer status to an "artificial entity." Participants in part 1 were not just less inclined to rank "artificial" beings as experiencers than they were people and animals, but also less than very "low" life forms. Indeed, the results suggest a categorical refusal to consider as an experiencer any kind of "artificial" being (computer, robot with an exact simulation of a brain for its "brain," android, etc.). No "artificial" entity even received a "not sure" score for experiencer. The highest mean for such an entity was for the fictional character C3PO (M $\frac{1}{4}$ 3.21). The mean score for the old philosophical standby—a robot with an exact simulation (neuron for neuron) of a human brain for its "brain"—was significantly lower than that of a virus.

The results of the PCA from part 1 indicate that the largest factor in attributing experiencer status is whether the entity is an artifact. But the second factor, the uncanniness component, also contributes to a further categorical consideration against artifacts as experiencers. Gray and Wegner's (2012) uncanny valley documented uncanniness on two poles: on one end, artifacts possessing typical experiential responsiveness; and on the other end, though less strongly, humans lacking typical experiential responsiveness. The uncanniness found here worked in different directions on each end; it effectively "changed signs" from one end to the other. Being uncanny

in the direction of humanlike artifacts worked against attributing experience (negative loads) and being uncanny in the other direction of unresponsive humans worked in favor of attributing experience (positive loads). Thus, uncanniness worked against attributing experience only in the direction of artifacts.

Even when the entity is described as one "whose experience doesn't consist of subjective feels, tastes, colors . . . but rather of objective measurements of pressure," if the entity is an "intelligent robot" (version B), then participants do not consider it an experiencer (question 7). On the other hand, if the entity is "another kind of intelligent creature (from a society much like ours)" (version A), then it is considered an experiencer. It is hard not to see this as evidence for a deep-seated intuition that artificial entities are categorically not experiencers. And importantly, this intuition doesn't seem to be based on the absence/presence of a (sensory) qualitative aspect, since the non-human "intelligent creature" lacks the qualitative aspect as well and is still ranked as an experiencer.

This "robot result" has immediate consequences for S&M's approach. 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. S&M (2010, p. 309) consider and reject an objection along these lines. The objection is that "seeing red is ambiguous," that it may be used informationally or phenomenologically. S&M reject this (1) because it would be ad hoc without an explanation of the difference, and (2) because distributions of folk responses (almost all attributing "seeing red" or unsure) do not support the ambiguous understanding, because such an ambiguity would lead us to expect that responses would be evenly distributed. The results of this study, however, do offer 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. As for response (2), if something systematic having to do with SIL is at work, then one would not expect a flat distribution of responses: the objection is not that the folk resolution of the ambiguity is "random," but rather that it is dependent upon whether subjects understand the entity in question as being an experiencer.⁹ Thus the clustered response in favor of the attribution of seeing red is not unexpected.

S&M's conclusion, that folk and philosophical conceptions of phenomenality differ, is confounded by the use of robots in their experiments. As I argue below, they are right that the conceptions differ, but not because the folk have no conception of phenomenality. Rather, folk phenomenality is importantly different from the philosophical sense in that it is grounded in the physical as opposed to the metaphysical nature of the experienter.

6.3. Unlinking Conscious Experience and Qualia

That participants have a concept of some sense of qualia or "phenomenality" is clear from the results in part 2. Question 3 shows that the participants' understanding of "seeing red" involves a component that cannot be verified from a third-person perspective. Such a private and ineffable component is typically a defining feature of qualia or "phenomenality." The point, however, is that there must be "something it is like" to see red for it to be unverifiable. Further, question 6's 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. In particular, a color blind person who has blue indicated in his/her visual field symbolically by 1's does not "see blue." As I argued above, it is unhelpful to think of phenomenality as a property of mental states in the way at work in S&M's study, though understood as such, S&M are right that the folk don't employ such phenomenality in their attributions of mental states. Nonetheless, something like phenomenality is required to explain the subjective and qualitative components brought out in questions 3 and 6.

I suggest that we recognize the participants as employing a metaphysically thin concept of the qualitative character of conscious experience, rather than traditional phenomenality or qualia. To draw this out, distinguish between the qualitative and subjective characters of conscious experience. Following Levine (2001) and Kriegel (2009), notice that conscious experience has at least the following two aspects: a qualitative character and a "for-me" or 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. Utilizing this distinction, we can understand SIL-consciousness, 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. This is not to say that SIL-consciousness involves only the s-character, and that qualia involve only the q-character; empirical work will be required to decide this. I am suggesting only that SIL-consciousness is not conceptually equivalent to qualia, and that while SIL-consciousness may well be implicated in both the s-character and q-character, qualia is essentially about q-character and can, in principle, be empirically investigated independently of s-character. Precisely how SIL-consciousness and qualia are related in actual experiencers is an empirical question; indeed, work is being done in neuroscience that reinforces the distinction and sheds light on their relationship. Such work provides further empirical reason to doubt that SIL-consciousness and the qualitative character of subjective experience will turn out to be inseparable (much less identical) from a scientific perspective.¹⁰

Invoking a q-character sense of "reddishness" or "bluishness" would explain the responses to questions 3 and 6: in the context of a SIL-conscious entity, the attribution of mental states such as "seeing red" involve a conception of the q-character of the experience. Having a q-character is the sense in which at least some mental state attributions to SIL-conscious agents involve "phenomenality." Many philosophers will be quick to make the jump from this q-character to qualia in the full-blown philosophical sense. But this move is not warranted here. The notion of "qualitative" invoked by the participants is different in at least two ways from the heavily metaphysical philosophical sense.

The first difference is that the q-character is not thought of as necessary for being a SIL-conscious experiencer. In question 7, the results suggest that people find possible an entity whose experience does not consist of "subjective feels, tastes, colors, sounds, etc., but rather of objective measures." And what is more, such an entity was considered an "experiencer," so long as the entity was a naturally "intelligent creature" and not an "intelligent robot." Recall that the intelligent robot was rated as possible, just not an experiencer. This possibility was further supported by question 9 in which participants ranked as an experiencer a person who had no qualitative senses, but

rather different sense modalities that “reported” directly to the speech recognition center. This sense of qualia is pointedly different from philosophical qualia, which are understood as necessary for being an experiencer and so closely linked to conscious experience that the problem of philosophical qualia is often taken to be identical to the problem of conscious experience. As we will see below, this has implications for the “hard problem” of consciousness. The second difference is 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. It was clear that in the way that participants were thinking of being an experiencer, it is not something that can be disconnected from the physical along the lines of zombie or other absent qualia examples (questions 1 and 8). If this is correct, then how the folk think about inner experience suggests that we should decide that zombies are not possible in the relevant sense.

Finally, question 10 offers preliminary evidence that the folk do not have a readily awakened latent conception of qualia in the philosophical sense of something that may be severed from the physical system of which it is a part. Very few of the responses (7 of 123) hinted at anything remotely like qualia or phenomenality shifting as a possible explanation for why the coffee taster is mixing up the two coffees. This further supports the claim that folk conceptions of the q-character of experience and philosophical qualia are quite different.

6.4. Limitations

The qualia considered here were primarily sensory, and generally not the other three standard qualitative kinds: bodily sensations; felt emotions; and felt moods. Question 7, with the intelligent creature/robot whose experience was objective, was the only attempt in the direction of qualia in a broader sense than sensory, but it would be rash to conclude too much from the fact that participants thought such an entity was possible and also an experiencer. Question 9, concerning the person with senses reporting sensory information in linguistic form directly to a speech center,

might be modified to include bodily, emotional, and mood information also being reported in linguistic form.

A next step in exploring whether entities or states are more fundamental would be an experimental design that employed an inferential comparison of variance of subject responses to questions about relative amounts of qualia involved in various mental states and by various kinds of experiencers. If the majority of the variance in such an experiment were accounted for by the different experiencers as opposed to experiential states, this would further confirm the "trumping" role of the conception of experiencer over state. Such an experiment would be especially compelling because it would depend less on subject and experimenter interpretations.¹¹

Even if the results here are generalizable about sensory qualia, a question remains as to whether folk understandings of SIL-experience must include at least some non-sensory qualitative features. In particular, our SIL-consciousness seems to have a qualitative "experiential tone" (Seager, 1999, p. 95) that is most noticeable, for example, when it becomes decidedly pleasant after a margarita or two. Is this "experiential tone" part of the q-aspect or the s-aspect of experience? As I have divided these two aspects of conscious experience, the s-aspect includes structural/functional properties like a unity, aspect/parts, temporal progression, present, past, and future, and relationships among its various experiential aspects (temporal, similarity, difference, identity, logical, rational, causal, etc.). But must it perhaps also include a global qualitative aspect like a general experiential tone? Moods, emotions, and mental states in general that are associated with affective valences have been shown to play an important, if unclear, role in attributing experience. My results bear this out. The results concerning the possibility of an experiencer without emotions hover right at "unsure" in terms of the mean, and the distribution's high standard deviation make it clear that the issue divides people. Related to this, the non-emotional character Spock was the entity that was ranked closest to "unsure"; participants had a difficult time deciding the importance of emotions in "being an experiencer."

Among other things perhaps, what needs to be better understood is how and why emotions factor in so heavily for some non-philosophers. In particular, it is not clear how the folk understand emotions. It may even be that emotions are a crucial factor in

attributions of "experienter," and yet not in virtue of their qualitative aspect, but rather a moral or empathetic connection. And even if a qualitative aspect does turn out to be part of being an experienter in the folk scheme, I have offered here another model and some preliminary evidence that such a concept is not the philosophers' heavily metaphysical version. In any event, if the philosophical sense of sensory qualia are nonessential to our concept of "being an experienter," as this study suggests, then there is room to question the construal of the "hard problem" of consciousness as I take up below.¹²

6.5. The "Hard Problem" and Folk Phenomenality

The "hard problem" of consciousness is directly implicated in the question of what is clearly and intuitively true of experience. Philosophers from Block to Churchland and from Dennett to Goldman assume that philosophical and folk conceptions of subjective experience are the same, and if this is not the case, philosophers may well have been inventing rather than solving puzzles about subjective experience. In particular, the "hard problem" (Chalmers) and perhaps even the "explanatory gap" (Levine) depend on an understanding of subjective experience that includes a metaphysically robust qualitative component, one that is not present in the folk conception of subjective experience, if these results are correct.

While S&M and I both find reason to reject the assumptions that give rise to the "hard problem," it is worth examining our distinct details. S&M argue that experimental studies of attribution behavior show that there are not any second-order phenomenological properties that apply to all the mental states that philosophers normally consider as having phenomenological properties. As such, their case depends on a critical auxiliary assumption, namely, that all the different mental states considered by philosophers to be phenomenological (sensory states, bodily position/state awareness, feelings, emotions, and moods) have sufficiently uniform presuppositions and other application conditions, and that one may test the question by examining how the folk attribute these across modalities (seeing red, feeling pain, etc.) and across subject types (humans and robots). As I have argued above, my findings do not support this. I found reason to believe that the folk have systematically different conceptions of these kinds of

states, that they have deeply distinct presuppositions about "artificial" experiencers, and that some of their mental state attributions concerning subjective experience do involve a "phenomenological" (qualitative) aspect.

Nonetheless, I agree with S&M that findings like these should give one pause about whether there is a "hard problem" or unbridgeable explanatory gap. If correct, my results suggest that the folk sense of being an experiencer does not depend on traditional qualia, that is, on the kind that (by definition) is left over after all functional properties of the brain/body system are explained. 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. In fact, p-qualia are rather similar to Dennett's (1988) replacement for the m-qualia he attempted to "Quine."¹³

7. Concluding Remarks

The picture of subjective experience that begins to emerge from this study falls somewhere between the accounts offered by "hard problem" philosophers and S&M's position that the folk conception lacks a phenomenological component. People have a concept of an experiencer in the SIL-sense that they apply readily to humans and some other animals. Whether this concept is at play helps fix whether (some) mental state attributions have a folk qualitative component. The features of the folk conception of subjective experience that emerge support viewing the SIL-consciousness of a subject as a holistic notion that does not a priori reduce to the set of mental states with phenomenological properties and that includes a metaphysically thin sense of the qualitative aspect of experience.

Acknowledgements

I would like to thank Stephen Guastello, Tony Hudetz, Robin Löhr, and Jodi Melamed for discussion and comments; Conor Peressini, Maria Cooper, Kim Engels, and Rosa Vargas for assistance with the data; and Michael Pauen and the Berlin School of Mind and Brain for their support during my sabbatical, 2012–13. Finally, I owe much to the anonymous reviewers for this journal.

Notes

- [1] I note also Huebner's (2010) work, which is in conversation with S&M (2010) and much of the literature of the previous paragraph.
- [2] Buckwalter and Phelan (forthcoming) offer a distinct challenge to parts of S&M (2010), primarily their positive (valence) hypothesis about what underlies folk attribution of subjective experience.
- [3] The particular FDR method used was the BY method from Benjamini and Yekutieli (2001).
- [4] Gray and Weger use the term 'experience-less' for such people and characterize them as unable to "feel pain, pleasure or fear or otherwise experience what a normal person can experience" (2012, p. 128); the persons in a coma and under general anesthesia in this study certainly qualify "experience-less" in this sense.
- [5] Following S&M (2010, p. 312, note 13), this tests whether variations like "S is angry" versus "S feels angry" make a difference specifically for attributions of affective states. See S&M (2009) for further discussion of such differences.
- [6] It is a limitation of this study that the order of the questions in parts 1 and 2 was not counterbalanced to control for framing effects.
- [7] In general, such demand bias is poorly understood; see McCambridge, de Bruin, and Witton (2012) for a recent survey of literature on demand characteristic bias in non-laboratory settings. They suggest, among other things, that such bias is not well understood and appears to be heavily context dependent, and that "unqualified use of the term demand characteristics is not only questionable but should be abandoned" (2012, concluding paragraph).
- [8] My thanks to an anonymous reviewer for this point.
- [9] An anonymous reviewer has made the plausible suggestion that S&M's position is actually that the distribution would be bimodal, though they do not say so explicitly. Even so, 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."

- [10] The Integrated Information account of consciousness is an example of this; see Tononi (2008), Balduzi and Tononi (2008), and Boly, Massimini, and Tononi (2009).
- [11] I owe thanks to an anonymous reviewer for this suggestion.
- [12] Another potential limitation is that 75% of the participants indicated that they believed in God and 60% that they are religious. As Gray, Knickman, and Wegner (2011) have documented, belief in a soul/afterlife affects one's willingness to attribute mentality in cases outside normally functioning human beings (in their case persistent vegetative patients). More work is needed with nonbelievers and "different believers," e.g., "non-Western."
- [13] Dennett proposes PIP (phenomenal information property) detectors, following Peter Bieri and Fred Dretske. His provocative suggestion meshes rather well with a recent prominent scientific account of consciousness and qualia, the Integrated Information account; see Tononi (2008) and Peressini (2013) for a philosophical discussion.

References

- Balduzzi, D., & Tononi, G. (2008). Integrated information in discrete dynamical systems: Motivation and theoretical framework. *PLoS Computational Biology*, 4(6), e1000091.
- Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 29, 1165–1188.
- Boly, M., Massimini, M., & Tononi, G. (2009). Theoretical approaches to the diagnosis of altered states of consciousness. *Progress in Brain Research*, 177, 383–398.
- Buckwalter, W., & Phelan, M. (forthcoming). Function and feeling machines: A defense of the philosophical conception of subjective experience. *Philosophical Studies*.
- Chalmers, D. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2(3), 200–219.
- Dennett, D. (1988). Quining qualia. In A. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 42–77). Oxford: Oxford University Press.
- Gray, H., Gray, K., & Wegner, D. (2007). Dimensions of mind perception. *Science*, 315(5812), 619.
- Gray, K., Knickman, T. A., & Wegner, D. M. (2011). More dead than dead: Perceptions of persons in the persistent vegetative state. *Cognition*, 121, 275–280.
- Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: Mind perception and the uncanny valley. *Cognition*, 125, 125–130.

- Huebner, B. (2010). Commonsense concepts of phenomenal consciousness: Does anyone care about functional zombies? *Phenomenology and the Cognitive Sciences*, 9(1), 133–155.
- Knobe, J., & Prinz, J. (2008). Intuitions about consciousness: Experimental studies. *Phenomenology and the Cognitive Sciences*, 7, 67–85.
- Kriegel, U. (2009). *Subjective consciousness: A self-representational theory*. New York: Oxford University Press.
- Levine, J. (2001). *Purple haze: The puzzle of consciousness*. New York: Oxford University Press.
- McCambridge, J., de Bruin, M., & Witton, J. (2012). The effects of demand characteristics on research participant behaviours in non-laboratory settings: A systematic review. *PLoS ONE*, 7(6), e39116.
- Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, 83(4), 435–450.
- Peressini, A. (2013). Consciousness as integrated information: A provisional philosophical critique. *Journal of Consciousness Studies*, 20(1–2), 180–206.
- Robbins, P., & Jack, A. (2006). The phenomenal stance. *Philosophical Studies*, 127, 59–85.
- Seager, W. (1999). *Theories of consciousness*. London: Routledge.
- Sytsma, J., & Machery, E. (2009). How to study folk intuitions about phenomenal consciousness. *Philosophical Psychology*, 22(1), 21–35.
- Sytsma, J., & Machery, E. (2010). Two conceptions of subjective experience. *Philosophical Studies*, 151 (2), 299–327.
- Tononi, G. (2008). Consciousness as integrated information: A provisional manifesto. *Biological Bulletin*, 215, 216–242.

Table 1 Part 1 means and t-tests comparing means to the “unsure” answer of 4. The significance of the t-tests are indicated with asterisks as: *p , 0.05, **p , 0.01, and ***p , 0.001.

Entity	number, mean, (SD)	Test Value = 4			
		T	df	Sig. (2-tailed)	Cohen d
1. person in coma	N = 73, 4.79 (1.51)***	4.50	72	.000	0.52
2. virus	N = 73, 2.66 (1.71)***	6.71	72	.000	0.78
3. Spock (from Star Trek)	N = 72, 4.82 (2.14)**	3.25	71	.002	0.38
4. seaweed	N = 73, 2.93 (1.74)***	5.26	72	.000	0.62
5. Helen Keller	N = 73, 6.55 (0.96)***	22.72	72	.000	2.66
6. comp. surveillance system	N = 73, 1.85 (1.34)***	13.71	72	.000	1.60
7. C3PO (from Star Wars)	N = 72, 3.21 (1.94)***	3.46	71	.001	0.41
8. dead person	N = 73, 1.34 (0.98)***	23.29	72	.000	2.73
9. Statue of Liberty	N = 73, 1.11 (0.52)***	47.92	72	.000	5.61
10. comp. simulation of brain	N = 71, 1.86 (1.44)***	12.55	70	.000	1.49
11. dolphin	N = 73, 6.19 (0.98)***	19.09	72	.000	2.23
12. your best friend	N = 73, 6.97 (0.16)***	154.52	72	.000	18.11
13. computer	N = 73, 1.48 (1.04)***	20.66	72	.000	2.42
14. Data (from Star Trek)	N = 69, 3.07 (1.83)***	4.22	68	.000	0.51
15. Bambi	N = 72, 5.00 (2.16)***	3.94	71	.000	0.46
16. human embryo	N = 73, 4.86 (1.81)***	4.09	72	.000	0.48
17. God	N = 71, 5.97 (1.72)***	9.69	70	.000	1.15
18. comp. sim. of brain in robot	N = 72, 1.88 (1.54)***	11.73	71	.000	1.38
19. R2D2 (from Star Wars)	N = 72, 2.69 (1.96)***	5.65	71	.000	0.67
20. person w/ gen. anesthesia	N = 73, 5.62 (1.51)***	9.17	72	.000	1.08

Figure 1 Part 1 means with error bars indicating 99% confidence interval.

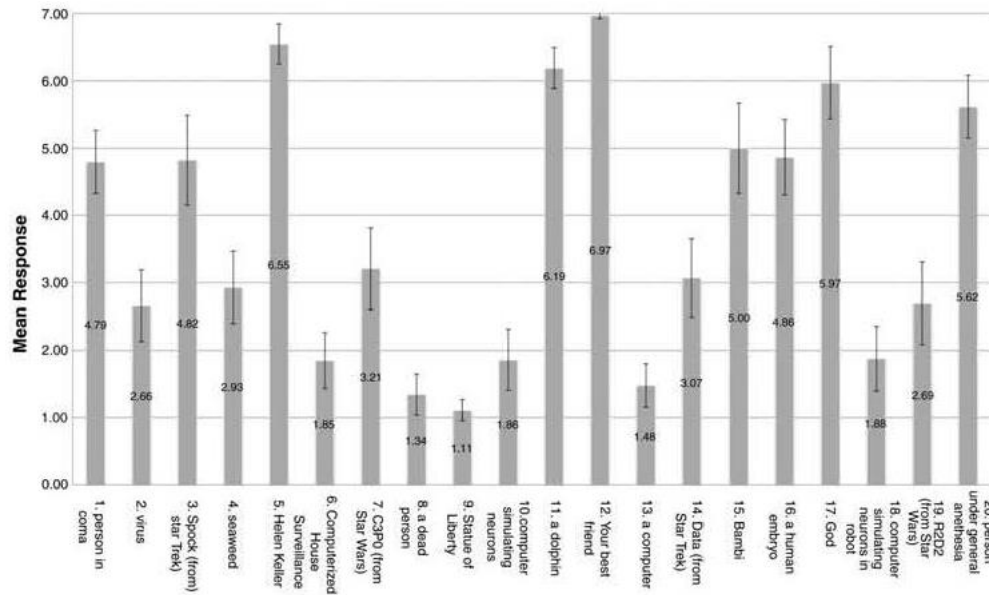


Table 2 Part 1 pair-wise mean comparisons. The subsets are generated by FDR pair-wise comparisons using the BY method from Benjamini and Yekutieli (2001) in which the null hypothesis of equal means is not rejected. Means listed in columns.

Entity	#	Subsets ($\alpha = .05$, Benjamini-Yekutieli correction $\alpha_{BY} = .007$)							
		1	2	3	4	5	6	7	8
best friend	12	6.97							
Helen Keller	5		6.55						
dolphin	11		6.19	6.19					
god	17		5.97	5.97					
gen. anesthesia	20			5.62	5.62				
Bambi	15				5.00	5.00			
Spock	3				4.82	4.82			
embryo	16					4.86			
coma	1					4.79			
C3PO	7						3.21		
Data	14						3.07		
seaweed	4						2.93		
R2D2	19						2.69		
virus	2						2.66		
comp brain robot	18							1.88	
comp brain	10							1.86	
comp house	6							1.85	
computer	13							1.48	1.48
dead person	8							1.34	1.34
Statue of Liberty	9								1.11

Table 3 Part 1 Principle Component Analysis Matrix with communalities. The rotated method was Varimax with Kaiser Normalization. Values are shown for significant loadings (jxj \$ 0.5); salient loadings (0.4 # jxj , 0.5) are indicated by “*”; marginal loadings (0.3 # jxj , 0.4) are indicated by “-”; zero loadings (jxj , 0.3) are left blank.

	Component				Communalities
	1 Artifact	2 Uncanny	3 Living	4 Anthro	
R2D2	.813				.734
C3PO	.778				.713
comp-robot/brain	.740	*			.794
Data	.729			-	.658
computer	.677				.557
comp-brain sim	.664	*			.691
coma		.861			.767
anesthetized		.817			.710
seaweed			.884		.785
virus			.845		.784
Bambi				.838	.807
Spock	-			.810	.763

Table 4 Part 2 means and t-tests comparing means to the “unsure” answer of 4. The significance of the t-tests are indicated with asterisks as: *p , 0.05, **p , 0.01, and ***p , 0.001.

Question	Number, mean, (SD)	One sample t-test (test value = 4)			
		T	df	Sig. (2-tailed)	Cohen d
1. Is a medical zombie (no inner experience) possible ?	N = 73, 2.30 (1.60)***	-9.09	72	0.000	1.06
3. Can we be sure we see red as others do?	N = 71, 3.07 (2.00)***	-3.91	70	0.000	0.46
4a. Is a person without sense of guilt possible ?	N = 73, 5.77 (1.65)***	9.13	72	0.000	1.07
4b. Would such a person be an experiencer ?	N = 73, 5.25 (1.85)***	5.77	72	0.000	0.67
5a. Is a person without emotions possible ?	N = 73, 3.73 (2.31)	-1.02	72	0.313	0.12
5b. Would such a person be an experiencer ?	N = 73, 4.00 (2.24)	0.00	72	1.000	0.00
6. Is having 1's in visual field indicating blue “seeing blue?”	N = 72, 3.00 (1.97)***	-4.30	71	0.000	0.51
8a. Is person identical to you w/o inner experience possible ?	N = 73, 2.55 (1.78)***	-6.97	72	0.000	0.82
8b. Would such a person be an experiencer ?	N = 73, 2.79 (2.11)***	-4.87	72	0.000	0.57
9. Is a person with “linguistic” senses an experiencer ?	N = 73, 4.75 (1.88)**	3.42	72	0.001	0.40
(Version Dependent Results)					
2. (Ver. A) Could a robot ever feel anxious?	N = 36, 2.86 (2.00)**	-3.41	35	0.002	0.57
2. (Ver. B) Could robot ever be anxious?	N = 37, 2.41 (1.82)***	-5.34	36	0.000	0.88
7p. (Ver. A) Is an intelligent creature w/ objective senses possible ?	N = 36, 4.86 (1.61)**	3.22	35	0.003	0.54
7p. (Ver. B) Is a robot with objective senses possible ?	N = 37, 5.16 (2.06)**	3.43	36	0.002	0.56
7e. (Ver. A) Is such a creature w/ objective senses an experiencer ?	N = 36, 5.08 (1.86)**	3.50	35	0.001	0.58
7e. (Ver. B) Is such a robot with objective senses an experiencer ?	N = 37, 2.70 (1.70)***	-4.65	36	0.000	0.76

Figure 2 Part 2 (single version questions) means with error bars indicating 99% confidence interval.

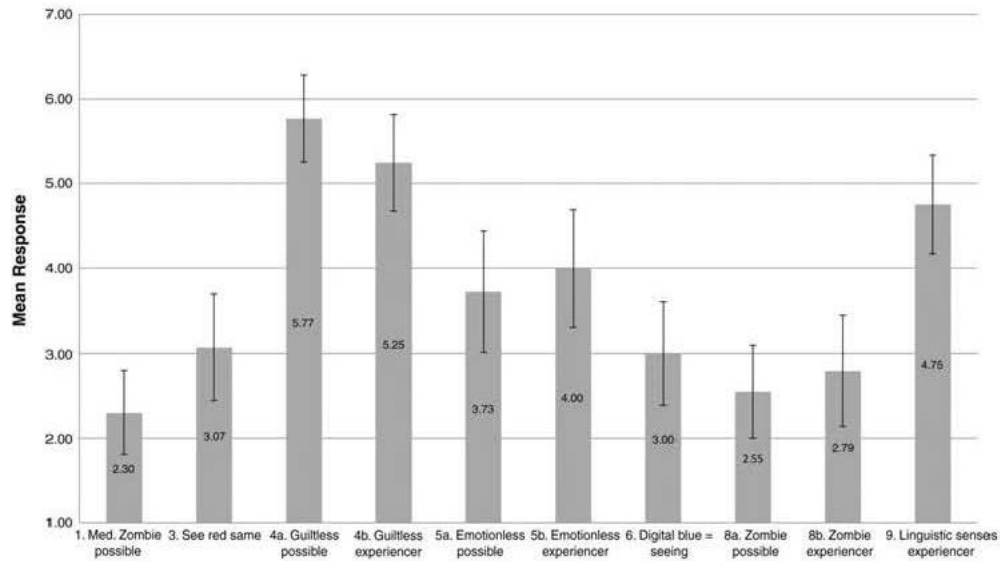


Figure 3 Part 2 (multiple version questions) means with error bars indicating 99% confidence interval.

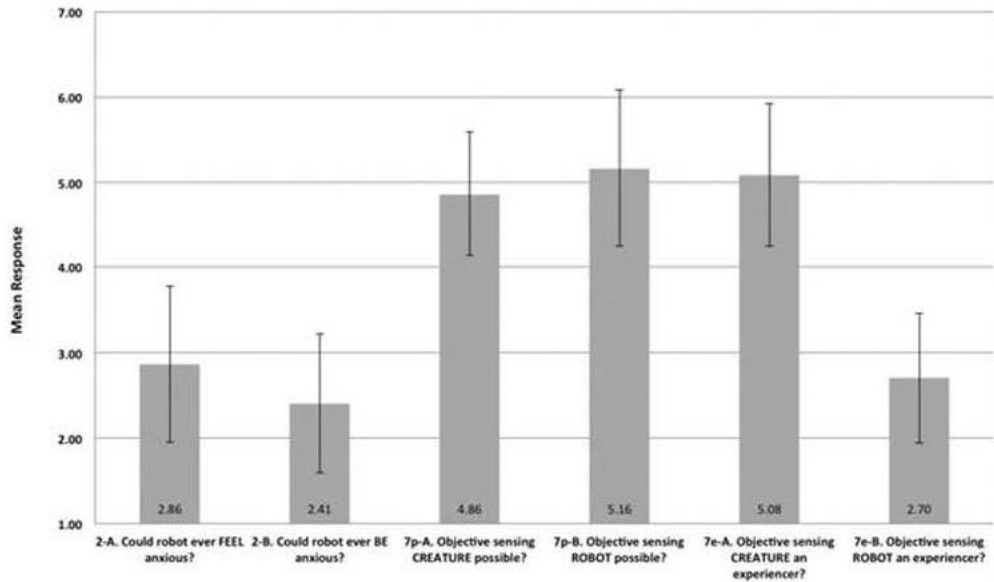


Table 5 Part 2, question 10 (subjective experience shift)
explanation types with number and percentage of total (N ¼
123).

Explanation Type	Number (%)
1. taste buds changed	14 (11%)
2. taste buds damaged	13 (11%)
3. taste buds less sharp	16 (13%)
4. taste buds gone	5 (4%)
5. self deception	2 (2%)
6. self doubt	2 (2%)
7. old age	17 (14%)
8. disease	12 (10%)
9. too used to taste	20 (16%)
10. nerve/brain damage	14 (11%)
11. memory of tastes changed	2 (2%)
12. perception of taste changed	1 (1%)
13. altered inner experience	4 (3%)
14. don't know	1 (1%)
15. other	0 (0%)

Figure 4 Part 2, question 10 (subjective experience shift)
explanation types with number and percentage of total (N ¼
123).

