

The Relevance of Folk Intuitions to Evaluating the Justification for the “Hard Problem”

Justin Sytsma and Edouard Machery

In his thought-provoking article, Brian Talbot offers an intriguing new objection to Sytsma and Machery (forthcoming). His objection is based on a novel account of how people make judgments about those mental states that philosophers typically take to be phenomenally conscious. We find this account quite interesting in its own right, and we are grateful for Talbot’s interesting contribution to the current research on the folk theory of consciousness. However, in this commentary, we will argue that Talbot’s main line of critique misses its target.

Here is how we will proceed. In Section 1, we show that as presented Talbot’s critique does not engage with the main argument that we gave in our article, due to a subtle mistake in his formulation of it. In a nutshell, we did not investigate folk intuitions to generate data directly about whether or not humans have phenomenal consciousness; rather, we turned to folk intuitions to test a common *justification* that has been given for thinking that phenomenal consciousness exists. Nonetheless, Talbot’s critique can be slightly modified to produce an argument that hits its mark. Thus, in Section 2, we offer a revised version of his objection and respond to it in Section 3.

1. The Argument

Here is how Talbot reconstructs the main argument in Sytsma and Machery (forthcoming):

Sytsma and Machery make the following argument: if we have phenomenal consciousness, then it is “the most central and manifest aspect of our mental lives.” This claim is an adoption of one made by those who think that there is a hard problem. If qualia or phenomenal consciousness are the most central and manifest aspect of our mental lives, then we should expect the folk to be aware of them. If the folk are aware of qualia, then this awareness should be reflected in folk judgments about mental states.

Specifically, it ought to play a role in how the folk ascribe mental states (or their lack) to various things in the world. However, Sytsma and Machery claim, it seems that it does not. So, by *modus tollens*, they conclude that there is no hard problem, or at least we have some good evidence that there is none. (2)

A key part of the framing of our argument is left out, however, leading Talbot to state a conclusion that we do not draw. Specifically, we do not conclude that there is no hard problem, nor do we conclude that we have good evidence that there is no hard problem. We do not claim to have produced *direct evidence* about the existence of phenomenal consciousness; rather, we claim to have produced evidence against a standard justification given for thinking that phenomenal consciousness exists (and, thus, *indirect evidence* about the existence of phenomenal consciousness).

As such, we do not rely on the first premise that Talbot gives. We do not claim that “if we have phenomenal consciousness, then it is ‘the most central and manifest aspect of our mental lives’” (quoting Chalmers, 1995, 207); instead, we note that the claim that phenomenal consciousness is the most central and manifest aspect of our mental lives has been given in support of the assertion that we have phenomenal consciousness. We challenge this claim, drawing the inference that Talbot notes: “If qualia or phenomenal consciousness are the most central and manifest aspect of our mental lives, then we should expect the folk to be aware of them.” Our empirical data is taken to show that the consequent is false and thus that the antecedent is false. We then conclude that the claim that phenomenal consciousness is the most central and manifest aspect of our mental lives does not support the assertion that we have phenomenal consciousness.

Once our argument is correctly articulated, Talbot’s objection does not clearly apply as given. He argues that “even if the folk do experience phenomenal consciousness, this should play no role in the sorts of judgments Sytsma and Machery... study” (3). This would be a

problem if we argued that these judgments constituted direct evidence that the folk do not have phenomenal consciousness; but, as noted above, that is not what we argue. Rather, we argue that these judgments are evidence that phenomenal consciousness is not a central and manifest aspect of our mental lives; or to put this more concisely, that phenomenal consciousness is not *obvious*.

2. Reformulating the Objection

Nonetheless, Talbot's objection can be naturally reformulated as an objection to our actual argument. He argues that the folk judgments we elicited are System One judgments and that System One judgments should not be expected to employ information about phenomenal consciousness (if such there be). System One judgments are typically generated quickly and unconsciously, based largely on associations; they are contrasted with System Two judgments, which are generated more slowly, consciously, and involve rational or rule-based thought (Talbot, 4; also see Stanovich and West, 2000).

Accepting the distinction between System One judgments and System Two judgments for the sake of argument, and assuming that if phenomenal consciousness is obvious that at least one of these systems will be sensitive to information about it, we can reformulate Talbot's objection as follows:

- (1) The folk judgments we elicited are System One judgments.
- (2) System One judgments, however, should not be expected to employ information about phenomenal consciousness.
- (3) As such, to test the claim that phenomenal consciousness is obvious we need to examine System Two judgments.
- (4) Therefore, we have not properly tested the claim that phenomenal consciousness is obvious.

This reformulated objection *is* an objection to the argument given in Sytsma and Machery (forthcoming). The objection concludes that we did not actually establish what we claimed to, as we examined judgments that could not be sensitive to the phenomenal qualities of mental states.

3. Responses to the Reformulated Objection

Each of the three premises of the reformulated version of Talbot's objection can be challenged, however. In this section, we will respond to each in turn.

3.1 Premise (1)

The first premise in the reformulated objection is problematic because it is not at all clear that the judgments we elicited are System One judgments. Talbot notes that we can suppress our System One judgments, reporting System Two judgments instead. In fact, there is a standard test in the literature to measure how likely a person is to suppress their System One judgments—Shane Frederick's (2005) Cognitive Reflection Test (CRT). As such, we can give participants one of our probes followed by the CRT and check whether or not CRT score correlates with the answers that participants give to the probe question. If System Two judgments diverge from System One judgments in the relevant way, then we would predict a negative correlation between participants' disposition to ascribe a mental state that is paradigmatically taken to be phenomenally conscious by philosophers and their CRT score. But that is not what we find.¹

297 Participants (72.1% females; age range: 18-77; age mean: 35.6) were presented with either the following vignette based on the one used in Sytsma and Machery (forthcoming) or a close variation on it that described Jimmy's pain behavior in more detail:

¹ The results reported were run as part of a larger study with Adam Arico.

Jimmy (shown below) is a relatively simple robot built at a state university. Jimmy is equipped with a video camera, wheels for moving about, and two grasping arms for moving objects. An array of sensors is embedded in the touch pads at the ends of the grasping arms.



As part of an experiment, Jimmy was put into a room that was empty except for one blue box, one red box, and one green box (the boxes were identical in all respects except color). Jimmy was instructed to put the blue box in front of the door. Jimmy did this with no noticeable difficulty. The test was then repeated on three consecutive days with the order of the boxes shuffled. On the first two days Jimmy performed the task with no noticeable difficulty. On the third day, however, when Jimmy grasped the blue box, Jimmy was given a strong electric shock! Jimmy immediately let go of the box and moved away from it. Jimmy did not try to move the box again.

Participants were then asked whether Jimmy saw blue (in addition to two other questions that we will not discuss here), on a 7-point scale anchored at 1 with “clearly no,” 4 with “not sure,” and 7 with “clearly yes.” Participants were also administered the Cognitive Reflection Test (mean CRT score: 0.54 out of a 4-point scale, where 0 is the minimum score and 3 the maximum, $SD=0.92$). By and large participants answered that Jimmy saw blue (mean response: 5.49) and there was no correlation between participants’ answers and their CRT scores ($r(297)=0.54$, $p>.35$). Because of the large sample size, it is unlikely that these results are due to a low power. Thus, contrary to what Talbot’s proposal suggests, people’s CRT does not predict how they ascribe a visual state like seeing blue to a simple robot.

It might be objected that even the high CRT participants did not use System Two judgments in responding to the test question, either because they were not aware of reasons to mistrust their System One judgments about this case or because they did not know how to make System Two judgments about this case. (In fact, Talbot himself suggests these replies.) But, if either explanation is true, then it is unlikely that participants’ judgments would have provided any evidence that ordinary people distinguish phenomenal states from non-phenomenal states, even if the judgments had been produced by System Two. If people spontaneously classify mental states as phenomenal and as non-phenomenal when System Two is engaged, then they

should know how to make the relevant judgments about this case, and at least high CRT participants should be able to produce reason to mistrust their System One judgments. We would expect people to spontaneously classify mental states in this way if the claim at issue—that phenomenal consciousness is obvious—is correct. That is, we would expect folk judgments of one type or the other to reflect this. If neither type of judgments are sensitive to information about phenomenal consciousness, however, then we again do not find it credible that phenomenal consciousness is obvious.

3.2 Premise (2)

The second premise in the reformulated version of Talbot's objection is also problematic:

Supposing that people have phenomenal consciousness, it is unclear why System One judgments would not generally employ information about it. Talbot's argument for this premise is based on the claim that System One judgments are associational. It is then argued that the relevant associations for ascriptions of non-valenced mental states will be between different third-person experiences with the entities ascribed such states, not between third-person experiences with those entities and our first-person experiences with the type of mental state being ascribed.

Talbot writes: "Thus, we will have much stronger associations between the external features of objects and their mental states than we do between data about qualia and mental states." (5). The key question is: Why shouldn't the associations be thought to be slightly more complex than this, involving associations between the behaviors of others, our own behaviors, and our first-person experiences with the type of mental state being ascribed? Talbot does not say much on this score. Here is what he says: "Now, we could form strong associations between qualia (if they exist) and mental states through first-person experiences of our own qualia and mental states. However, this is likely to play relatively little role in System One judgments about *others'* mental states."

(5). It nonetheless remains unclear to us why our first-person experiences with the mental states being ascribed should be expected to play relatively little role in the ascriptions.

Note that the associations at issue cannot just be between the external features of different entities (call these *behaviors*), but must also be involve mental states. How do the associations between behaviors and mental states get made? This is a difficult and complex empirical question, but it is possible that associations between *one's own* behaviors and *one's own* mental states will play *some* role in this. If this is correct, then the relevant associations will be between the behaviors of others', one's own behaviors, and one's own mental states. It would then be unclear why System One judgments would not employ information about phenomenal consciousness if it is obvious in having the relevant mental states that they are phenomenally conscious.

In addition, we are not really convinced that Talbot's model of mental state ascription is plausible, as we ascribe mental states to creatures that are very different from typical humans and that do not exhibit all of the behaviors that humans do. We can ascribe sensory experiences to a newborn, who is barely able to move; people ascribe mental states to all kinds of animals; Japanese children ascribe mental states like pain to *tulips* (Inagaki & Hatano (1987, 1991); Arico and colleagues (ms) have provided evidence that Westerners have an initial tendency to ascribe mental states like pain to *insects* and to *plants*; and so on and so forth. Perhaps Talbot would reply that such ascriptions are done by System Two, not System One, but this reply is dubious: for instance, Arico and colleagues insist that the ascriptions of pain to insects and plants results from an automatic, defeasible process.

What's more, it is not clear at first glance that Talbot's account of mental state ascription is well-suited to explaining the overall body of data reported in our original article. Thus, he

gives a *further* explanation for why the relevant associations for ascriptions of valenced mental states would include first-person experiences with those mental states:

When we encounter others that are experiencing valenced mental states – pleasure or pain, or anger or joy, etc. – we generally experience more than their behavior or appearance. We have a set of mental systems that often cause us to simulate the experience of valenced or emotionally charged mental states when we see another who is experiencing that mental state. (5)

More specifically, behavioral cues cause us to simulate valenced mental states, which involves having first-person experience with the corresponding mental states, so System One judgments about valenced mental states will involve associations between behavioral cues and first-person experiences. People do not judge that the robot felt pain, then, because it in some way fails to elicit the appropriate first-person experience that is needed for the association. For the sake of argument, we will not challenge the role of simulation in ascribing states like pain. What matters for present purposes is that to explain the role of valence, Talbot is compelled to assume that people's System One judgments are sensitive to their experiences—viz. to the experiences they have when they simulate the mental states that others might have. But, if this is the case, then it is unclear why something similar couldn't be going on when people ascribe mental states more generally. Accepting that people's System One judgments are sometimes sensitive to their experiences, why shouldn't we think that System One judgments are also sensitive to their experiences for ascriptions like seeing red?

3.3 Premise (3)

The inference of premise (3) from premise (2) in the reformulated version of Talbot's argument is also problematic. Even if it were the case that System One judgments cannot be sensitive to the distinction between phenomenal and non-phenomenal states (a claim we just challenged) and are thus useless to test philosophers' claim about the obviousness of phenomenal consciousness,

it does not seem that System Two judgments are the right sort of judgments to support the claim that phenomenal consciousness is obvious. It is not clear exactly what is meant by claims that phenomenal consciousness is the most central and manifest aspect of our mental lives, for example, but if judgments that it is so depend on conscious, rule-based thought (i.e., System Two processes), then something more should be said: What is it about the deliberative thought that renders phenomenal consciousness obvious? Talbot's proposal seems to suggest that phenomenal consciousness would not be obvious just in having the relevant experiences, but that further considerations would be involved in finding it to be obvious.

To illustrate, consider one of the CRT problems: "A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost?" The incorrect answer of \$0.10 tends to spontaneously jump out at people, although some quick checking can make clear that the answer is incorrect. This leads some people to think further about the problem. When they do so, they tend to arrive at the correct answer (\$0.05). While it is natural to say that the initial answer seems obvious to people, we might also say that the correct answer is obvious given some further thought. In this case, however, the latter obviousness claim can be readily supported with a further clause: "The answer of \$0.05 is obvious because then the bat would cost \$1.05 and \$1.05 plus \$0.05 is \$1.10." If this analogy is apt, and if phenomenal consciousness is thought to be obvious in a similar way, then it is fair to ask for the further clause explaining what renders phenomenal consciousness obvious. Given such an explanation we could then consider whether it is as compelling as the explanation given for the CRT question is, or whether it instead involves assumptions that can be readily questioned.

In fact, there is reason to believe that System Two judgments involve one's theory about the domain at issue and, as such, that for purposes of establishing that phenomenal consciousness is obvious one should appeal to the judgments made by System Two only with trepidation.

Although there is a fair amount of disagreement among psychologists about the nature of System Two, System Two judgments seem to be often influenced by whatever folk theory people have been taught (see, e.g., Haidt's [2001] on reflective moral judgments). One should thus be concerned that instead of reflecting the properties of their experiences System Two judgments might reflect whatever folk theory of subjective experience is prevalent in a given culture. With regard to philosophers, we should then be concerned that their System Two judgments reflect their philosophical theory of consciousness. But, that phenomenal consciousness is obvious given a particular philosophical theory does not carry the same weight as does the claim that phenomenal consciousness is pretheoretically manifest (as we discuss in Section 6 of our original article).

4. Conclusion

We have argued that Talbot makes a subtle mistake in formulating the argument given in Sytsma and Machery (forthcoming). Correctly formulated it is not clear that Talbot's critique applies. Nonetheless, his critique can be reformulated so that it does apply to our argument. We then showed how this might work and offered several responses to the reformulated objection. Even when reformulated, however, we find that the objection is not convincing.

References

Arico, A., Fiala, B., Goldberg, R. F., and Nichols, S. Ms. The Folk Psychology of Consciousness.

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, 19(4): 25–42.

Inagaki, K. and Hatano, G. (1987). Young children's spontaneous personification as analogy. *Child Development*, 58, 1013-1020.

Inagaki, K. and Hatano, G. (1991). Constrained person analogy in young children's biological inference. *Cognitive development* 6: 219-31.

Stanovich, K. E. and R. F. West (2000). "Individual differences in reasoning: Implications for the rationality debate?" *Behavioral and Brain Sciences*, 23: 645–726.

Sytsma, Justin and Edouard Machery (forthcoming). "Two Conceptions of Subjective Experience." *Philosophical Studies*.

Talbot, Brian (manuscript). "The Irrelevance of Folk Intuitions to the 'Hard Problem' of Consciousness."