The Irrelevance of Folk Intuitions to the "Hard Problem" of Consciousness
Forthcoming in *Consciousness and Cognition*

Brian Talbot
Brian.Talbot@Colorado.edu
University of Colorado, Boulder

Abstract: Recently, a number of philosophers have turned to folk intuitions about mental states for data about qualia and phenomenal consciousness. In this paper I argue that current research along these lines does not tell us about these subjects. I focus on a series of studies, performed by Justin Sytsma and Edouard Machery, to make my argument. Folk judgments studied by these researchers are mostly likely generated by a certain cognitive system – System One – that will generate the same data whether or not we experience phenomenal consciousness. This is a problem for a range of current experimental philosophy research into consciousness or our concept of it. If experimental philosophy is to shed light into phenomenal consciousness, it needs to be better founded in an understanding of how we make judgments.

1. Introduction

Certain mental states, such as perceptual ones (seeing red), emotions (feeling anger), and bodily sensations (pain), are supposed by many philosophers to have a qualitative aspect – a "what it is like" to be in those mental states. The experience of the qualitative aspect to these mental states – of qualia – is phenomenal consciousness, and explaining the existence of qualia and phenomenal consciousness is what David Chalmers (1996) calls "the hard problem" of consciousness. Recently, there have been a number of papers that look to folk judgments (i.e., judgments of non-philosophers) for data relevant to the debate on the hard problem of consciousness (e.g. Knobe & Prinz, 2008, Sytsma & Machery, 2010, Huebner, 2010). This approach should seem at least initially reasonable, since philosophers and lay people alike are intimately acquainted with consciousness. However, as I will argue in this paper, folk judgments as currently studied do not shed light on the existence of qualia or phenomenal consciousness, our experience of them, or the role they may play in our concepts of mental states.

My discussion will focus on one particular series of experiments on folk intuitions carried out by Justin Sytsma and Edouard Machery (2010). This is because while their research is among the strongest experimental philosophy work so far on this topic, it has in common with this other work the fact that it studies folk judgments about the mental states of other beings – beings other than the participant in the experiment. By showing how this by itself undermines the usefulness of the data Sytsma and Machery have gathered, I will illustrate the general problem for this line of research.

I will start by first giving the argument Sytsma and Machery make, and describing the data they generated to support this argument. I will then argue that the cognitive system that likely generates the judgments studied by Sytsma and Machery, and other experimental philosophers in this area, is not equipped to give useful data on phenomenal consciousness. At the end, I will generalize from my discussion of Sytsma and Machery's work. I do not take my arguments to be against the use of folk intuitions generally in philosophy, nor against the possibility of folk judgments, properly gathered and used, being a good source of data about consciousness, and at the end of the paper I will briefly discuss some ways one might do research so as to avoid the problems I have raised.

2. Sytsma and Machery's Argument

Sytsma and Machery's paper is concerned with the question of whether or not qualia, and phenomenal consciousness, are pre-theoretically apparent to the folk. This is of interest to them because the philosophical debate on the hard problem of consciousness typically begins from the premise that there *is* a hard problem at all, that qualia and phenomenal consciousness are there to be explained. This premise is, according to Sytsma and Machery, justified by appeals to our pre-theoretical acquaintance with phenomenal consciousness; phenomenal consciousness is, according to Chalmers, **"**the most central and manifest aspect of our mental lives."[1] Sytsma and Machery make the point that, 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. Thus, answering the question of whether or not the folk are aware of qualia is vital for understanding how seriously we should take the hard problem of consciousness.

How are we to answer this question? Sytsma and Machery argue that, if the folk are aware of qualia, then this awareness should be manifested in folk judgments about mental states. Specifically, it ought to be reflected in how the folk ascribe mental states, or their lack, to various things in the world. After reporting the converging results of a number of subtle and well-designed studies, they argue that this awareness seems not to be reflected in folk judgments, which suggests that the folk are not aware of qualia, and thus that qualia are not obvious. This is a serious blow to the assumption upon which the hard problem is founded.

I take issue with Sytsma and Machery's expectation that folk awareness of qualia should be reflected in the type of judgments that they studied. As I will show, given what we know about the mind and how it generates judgments, we should expect to find the results Sytsma and Machery obtained whether or not the folk are aware of qualia and phenomenal consciousness. Before I make this argument, let me first explain in more detail the data Sytsma and Machery obtained and how they obtained it.

3. Sytsma and Machery's Studies

Sytsma and Machery performed a series of studies in which they exposed participants to vignettes describing the reactions of either a human or a simple robot to various stimuli. They then asked participants if the being described (the human or robot) experienced a specific mental state – e.g. "Did Jimmy see red?" – and they took participants' answers to indicate whether or not they ascribed that mental state to that being.[2] Some of these mental states would be classified as qualitative by philosophers, and others not. They learned that folk participants are willing to ascribe certain types of supposedly qualitative mental states to very simple robots. However, participants were not willing to ascribe other supposedly qualitative mental states to these beings. The folk reported that a robot could see red and smell certain smells (that of Isoamyl Acetate), but could not feel pain or smell other smells, such as that of a banana.[3]

---

[1] Sytsma & Machery (2010), citing Chalmers, 1995, 207.

[2] I will not dispute that the gathered responses did track participants' mental state ascriptions.

[3] Subjects in each experiment were presented with some variation of the following stimulus: "Jimmy (shown below) is a relatively simple robot built at a state university. He has a video camera for eyes, wheels for moving about, and two grasping arms with touch sensors that he can move objects with. As part of a psychological experiment, *he was put in a room that was empty except for one blue box, one*

This data should seem inconsistent with folk awareness of qualia. Prototypically conscious states were ascribed to beings that do not have phenomenal consciousness. What's more, the supposed qualitativeness of mental states seemed to play no role whatsoever in participants' ascriptions of these states to these beings, since some qualitative mental states but not others were attributed to robots. There is no obvious, straightforward explanation of this data available to those sympathetic to the existence of qualia. For example, it cannot be attributed to robots' judged lack of the appropriate perceptual faculties, since they were judged able to smell some smells but not others. Nor can it be attributed to the participants taking the robots to be conscious, since the robots were judged *un*able to smell certain things, which makes no sense if robots were conscious and able to smell. It also seems unlikely these participants interpreted words like "see" or "smell" differently than do philosophers (e.g. using them to refer to purely functional states), since this does not explain why they reported that robots could smell some things but not others, or could see colors but not smell bananas.

Sytsma and Machery's claim is that these judgments track the distinction between valenced and non-valenced mental states, and not that between qualitative and non-qualitative mental states, or that between conscious and non-conscious beings. Valenced mental states are those that are associated with goodness or badness for the agent in question; often they are pleasurable or painful mental states. Non-valenced mental states are not associated with goodness or badness. For example, seeing red, or smelling Isoamyl Acetate is (to participants at least) non-valenced, but smelling a banana or feeling pain are valenced. Participants attributed apparently non-valenced mental states to robots, but not valenced ones. Sytsma and Machery argue that the valenced/non-valenced distinction is independent of the qualitative/non-qualitative distinction, and that the role of valence in these judgments does not generate a hard problem of consciousness. Since participants' judgments were based on valence and not qualitativeness, they argue, these judgments do not manifest an awareness of phenomenal consciousness, which in turn suggests that the folk are not aware of qualia or phenomenal consciousness, and undermines a core justification of the hard problem of consciousness.

In what follows, I will not question Sytsma and Machery's data. I will also not dispute that the difference in ascriptions of mental states to robots was caused by the valence, or lack thereof, of the mental state in question, nor will I dispute the claim that the valence of a mental state is independent of its being qualitative. Instead, I will argue that, even if folk judgments of others' mental states do track the valenced/non-valenced distinction, and this distinction has nothing do with qualia or phenomenal consciousness, we are not justified in concluding that the folk are not aware of phenomenal consciousness or qualia. Once we understand how the sorts of judgments Sytsma and Machery studied are produced, we will see that we should expect the results they obtained whether or not qualia are obvious. To see why this is, we must now discuss how our minds produce the sorts of judgments Sytsma and Machery studied.

---

red box, and one green box (the boxes were identical in all respects except color). An instruction was then transmitted to Jimmy. It read: '*Put the red box in front of the door*.' Jimmy did this with no noticeable difficulty. *Did Jimmy see red*?" (Sytsma & Machery, 2010, emphasis mine). The emphasized portions were changed from study to study or condition to condition to reflect the different qualitative states being tested (smelling banana, etc.) The pictured robot was very simple and bore no resemblance to a human.

4. System One Judgments about Mental States

4.1. What is System One, and How Does it Work?

We make judgments in a number of ways. Consider the judgment "Cats are animals." We can easily think of a number of different ways this judgment might come about. One might judge this to be true because it simply strikes him or her as true. One might also judge this to be true because one recalls having been told that it is true by a trusted source. One also might judge it is true because one has consciously recalled a number of traits cats share, and has compared these traits to those one believes animals must possess. Each of these ways of making this judgment draws upon cognitive systems that are, to some extent, distinct, both in terms of what information they draw upon (e.g. testimony about, versus experience of, cats) and how they use that information. Sytsma and Machery are attempting to do some "reverse engineering" - they are attempting to determine what information the folk are aware of by considering the judgments the folk make. Proper reverse engineering of this sort requires knowing not only what judgment has been produced, but also what cognitive system produced it, and how that cognitive system produces judgments.

My view is that the judgments Sytsma and Machery studied were a sort of intuitive, snap-judgment (I will argue for this below), and so we must understand how the cognitive system that produces these judgments does so before we can learn anything from their data. The snap-judgment producing system draws upon some different sources of information to make judgments than does our conscious reasoning system. If, for example, I have just been taught the rules to a complex game, I may be able to consciously draw upon these rules to determine what moves in the game are legal, but it is unlikely that my snap-judgment generating system will be able to use these rules. However, if I have played the game for years, I will at some point have automatic, intuitive judgments about what moves I can and cannot make. Conversely, the snap-judgment, or intuition, generating system can also draw upon information that we do not have conscious access to (see, e.g., Lewicki, et al, 1992) – for example, we all make automatic judgments about the grammaticality of sentences in our native languages, even when we do not consciously know the rules that make these sentences (un)grammatical.

In what follows, I will call the cognitive system that does (relatively) slow, conscious, and largely rational thinking "System Two," and I will use the name "System One" to refer to the intuitive, snap-judgment generating system. These names are borrowed from the psychologists Keith Stanovich and Richard West (2000). System One is faster than System Two, does most of its work unconsciously, and makes judgments based largely on associations.[4] I will argue in a minute that System One is largely responsible for the folk judgments Sytsma and Machery (and by extension, other experimental philosophers in this area) study. But first let us establish why this matters: it matters because System One will produce the same judgments about others' mental states whether or not we are aware of qualia or phenomenal consciousness, which means that we should not look to System One judgments about others' mental states to determine whether or not we are aware of qualia. Let us see why that is.

---

[4] For more on Systems One and Two, see Sloman, 1996, or Kahneman & Frederick, 2002. The discussion of the functioning of System One that follows is based on a large body of research that is explored in more depth in Talbot (2009, How to Use intuitions in Philosophy, University of Southern California, unpublished dissertation).

System One judgments are associative and similarity based: judgments about a current object are based on its similarity to things we have encountered in the past. When System One judges Fido to be a nice dog, it is because Fido bears a strong resemblance to Tagger, Lily, and Harley, all of which we have previously judged to be nice dogs, or because Fido is in a context (say, at a dog park) in which we have encountered other nice dogs in the past. The stronger the resemblance between Fido and past nice dogs, or the more often nice dogs have been found in dog parks, the more likely System One is to judge Fido to be a nice dog. So, if Fido looks and acts just like Tagger, then System One is going to tell me that Fido is a nice dog. Alternately, if Fido resembles Harley only somewhat, but I am in a dog park, then System One is likely to judge that Fido is a nice dog.[5]

It should be unsurprising that experience, both present and past, plays a huge role in System One's judgments. Our senses almost never stop working and giving us information, and so, generally speaking, most of the information we have about past objects comes from experience of those object. Similarities whose expression can be and is directly experienced will thus often sway System One more than similarities that are not as directly experienced. This is born out in ordinary life. If Fido is snarling and foaming at the mouth, experience tells us that Fido is not a nice dog. Imagine now that Fido's owner tells us that that is just how Fido gets when his teeth are being brushed, and that really Fido is very nice. Presumably we have also been told in the past that a certain nice dog was a nice dog. So, once we have been told that Fido is nice, Fido bears some similarity to nice dogs, and some similarity to bad dogs. However, even if we have encountered relatively few bad dogs in the past, there are a large number of similarities between Fido and these dogs, since Fido looks and sounds like them in a number of ways. On the other hand, there is only a small similarity between Fido and good dogs. And so our gut tells us that Fido is a bad dog, even if our conscious mind can reason to the conclusion that he is not.

4.2. How Does System One Generate Judgments About Others' Mental States?

Let us apply this to System One judgments about mental states. When System One ascribes a mental state to some object, it looks for similarities between that object and past objects that had that mental state. These will mostly be traits that we experienced those past objects as having. In other words, to decide if some being is seeing red, System One compares facts about this being to experienced facts about past objects that have been judged to see red. What sorts of similarities are likely to be employed by System One when making associative judgments about others' mental states?[6]

The discussion in the previous section should make it clear that the external features and behavior of beings should play a large role in these judgments. In deciding if Wall-E feels pain, System One will consider if Wall-E looks and acts like beings in the past that felt pain. We might expect these external features to play the dominant role in

---

[5] Of course, I cannot hope to exhaustively explain an entire cognitive system in one paragraph, but the nuances of the system I have just described are irrelevant to the arguments I am about to make.

[6] In what follows, I will contrast the judgments we should expect from System One if there were or were not qualia. What I will say applies equally well to what we should expect if we were or were not aware of qualia, which is what Sytsma and Machery are most immediately concerned with.

such judgments, but I do not want to assume this too quickly. What other traits could System One look to to find similarities upon which to ascribe mental states?

It should be obvious that others' qualia, if they exist, cannot play a direct role in System One judgments about others' mental states. Since we necessarily never experience others' qualia, System One can never consider whether the qualia a given being actually is or is not experiencing are similar to those experienced by beings in the past who were in a given mental state. But can qualia, if they exist, play an indirect role in System One judgments about others' mental states?

From time to time we might explicitly think about others' qualia when also thinking about their mental states, and this might provide some basis of similarity upon which to ascribe mental states. For example, we might witness someone stubbing their toe and intentionally imagine what they are going through when that happens; this would provide us with some data upon which future, automatic, pain attributions could be based. But the number of times we do intentionally imagine others' mental states is vanishingly small compared to the number of times we do not, but only perceive their external traits. We should not expect such explicit, intentional cognition to make much of a difference to System One judgments about mental states.

There are, however, more automatic mental processes that do something akin to what we just considered. When we see another person's facial expression, we often unconsciously and automatically simulate making that expression in our own minds.[7] The relationship between mental states and facial expressions is a two-way street: having these facial expressions is not only caused by the associated mental state, it also causes it (to a much lesser degree). Smiling actually makes one a little happier. This happens even when we simulate these expressions internally. If we have phenomenal consciousness, then when we successfully simulate another's qualitative mental state, we will experience the same sort of qualia they are experiencing. This is something System One might use in ascribing mental states to beings. Say we see Wall-E, for example (the robot from the eponymous Pixar movie), who has a sad look in his robot eyes. Seeing Wall-E's expression causes us to simulate that expression in our own heads, generating some sadness qualia. When System One is considering whether or not to ascribe sadness to Wall-E, our current experience of sadness qualia when seeing Wall-E would be a point of similarity between Wall-E and past sad beings. Conversely, if we did not experience sadness qualia when looking at the Terminator, because the Terminator is expressionless, we would be less willing to ascribe sadness to it, since this would be a dissimilarity between it and beings we had judged to be sad in the past. In this way, qualia would be able to play a role in System One ascriptions of mental states.

Putting this all together, we learn that whether or not qualia exist, System One will be willing to ascribe some allegedly qualitative mental states to beings that had no phenomenal consciousness, as long as those beings were externally similar enough to other beings that we took to have those mental states. First, consider mental states whose experience is not generally reflected in our facial expressions. Perceiving another in such

---

[7] For an overview of the literature on the material discussed in this paragraph, as well as a more detailed analysis of the evidence for and mechanisms behind these processes, see Goldman & Sripada, 2005. For the sake of simplicity I have left out some nuances of the different accounts of just how this works; for example, some theorists think that we simulate emotions *first* and then the appropriate facial expressions. The effect is the same for my purposes.

a mental state does not cause us to automatically simulate that states. So, relevant qualia would rarely be experienced when making judgments about others' mental states. We should expect System One ascriptions of these mental states to be based almost entirely on external features. What external features are commonly possessed by beings that are in mental states such as these (e.g., seeing red)? These will be things like having the appropriate receptors (eye-like things), being oriented towards appropriate objects (red things), and behaving appropriately. Since beings with no phenomenal consciousness can possess all of these, we should expect mental states that do not have associated expressions, such as seeing red, to be ascribed to some robots by System One, whether or not there is such a thing as phenomenal consciousness.

When we consider System One judgments about mental states that have associated facial expressions, we get somewhat different results. We have an additional, important basis upon which System One can make similarity judgments about these states: facial expressions, and possibly the qualia that simulations of them generate in us. And so we should not expect System One to attribute such mental states to beings that do not have the appropriate sorts of faces. This probably explains why Pixar, for example, went to so much trouble to give Wall-E mobile and expressive eyes, and why Arnold Schwarzenegger, with his relatively impassive face, was a good choice to play the Terminator.[8] If qualia do exist, they do play some role in ascriptions of these sorts of mental states. However, System One will generate the same judgments about such mental states whether or not qualia exist. We do not simulate others' mental states without provocation, but rather in reaction to their features. System One takes some input – external features of Bruce and his environment – and produces some output – a judgment about Bruce's anger. If there are qualia, then qualia play a role in the route from input to output. But in a world without qualia, we would see Bruce's features, simulate the appropriate mental state, experience no qualia, and judge that he is angry; after all, if we never experienced qualia when seeing angry people, then not experiencing it when seeing Bruce would not affect System One's judgment.

Valenced mental states, such as feeling pain, or experiencing pleasurable sensations, tend to be reflected in our facial expressions, but non-valenced mental states like seeing red or smelling Isoamyl Acetate do not.[9] So, we should expect System One to be willing to attribute non-valenced, but supposedly qualitative, mental states to the robots described in Sytsma and Machery's studies, because they can behave in a certain manner. And we should expect System One to not attribute valenced mental states to these robots, because they lack the appropriate facial features. Sytsma and Machery's results are thus to be expected if their participants were largely reporting System One judgments. Crucially, these results are to be expected from System One whether or not qualia exist, and whether or not the participants were aware of qualia. If Sytsma and

---

[8] The region immediately around our eyes conveys a massive amount of information about our emotional states – see, e.g. Baron-Cohen, et al (2001) – and Wall-E can convey emotion despite lacking most other facial features.

[9] This does not require that every valenced mental state have a corresponding facial expression, or be accompanied by some emotion-signaling facial expression. Associations exhibit some degree of transitivity (Lewicki, et al, 1994). So, if we have strong associations between many valenced mental states and characteristic facial expressions, and between these mental states and other valenced mental states without characteristic facial expressions, then how we attribute one mental state will affect how we attribute the other.

Machery's participants were largely reporting System One judgments, their data does not give us any insight into the hard problem of consciousness.

5. Why System One is Responsible for Sytsma and Machery's Data

The above would be a problem for Sytsma and Machery's argument if one of two conditions obtain: they think that System One judgments are especially important to appeal to in making the argument they wish to make, or if there are good reasons to think that the judgments they elicited were largely generated by System One. Sytsma has suggested that the first is the case.[10] But perhaps they need not appeal to System One judgments to make their argument. The question then is: do we have reason to think that, in any case, they are?

System One judgments are automatic and effortless (Kahneman & Fredrick, 2002). This means that they occur whether we want them to or not, even when we are not interested in them: "The associative system... always has its opinion heard..." (Sloman, 1996, 15) This does not mean that we always use or report our System One judgments. We can suppress them or replace them with judgments made in other ways. But this only happens when we (a) are aware of reasons to mistrust a given System One judgment, (b) care about getting the right answer in the case at hand, and (c) know of what we consider to be a better way of making the judgment in question. Otherwise, people tend to simply make and report System One judgments (see, e.g. Kahneman & Fredrick, 2002, Sloman, 1996). If we have no reason to think that the participants in Sytsma and Machery's study met conditions a, b, and c, we should take their judgments to be System One judgments.

The folk participants Sytsma and Machery studied likely did not meet these three conditions. They were in what seemed to be a low stakes situation, a psychological experiment about a seemingly easy question, their answer to which almost no one will know, and which will affect no one as far as they know. There was no obvious reason to mistrust their gut judgments on the matter, and no clearly better way to approach the question.

The above claim may seem inconsistent with the argument I have been making, however. If what I have been saying is true, and if there are qualia, then System One should often incorrectly ascribe qualitative mental states to beings; for example, it will tell us that cameras "see red." Would not this cause us to learn that System One judgments about others' mental states are not to be trusted? I do not think it would. Learning to mistrust System One requires not only that System One often makes mistakes, it also requires that we become aware of these mistakes, that something calls our attention to the fact that System One has judged incorrectly. This is unlikely to happen very often for System One judgments about mental states, because acting on the mistakes System One makes will have little cost, and thus these mistakes will not be called to attention. Unreflectively acting as if cameras see red, for example, will not cause the sort of undue distress that brings about reflection. Even when we do notice that these gut judgments are odd, they typically would not be problematic enough for us to go to the trouble to figure out that they are *wrong*. Of course, each of these is somewhat

[10] This is suggested by some of Sytsma's comments during the 2010 Online Consciousness Conference; see http://consciousnessonline.wordpress.com/2010/02/19/the-irrelevance-of-folk-intuitions-to-the-%E2%80%98hard-problem%E2%80%99-of-consciousness/

variable; some people are more inclined to second-guess their answers, or to take experiments more seriously. And that is reflected in the degree of variation in responses to some of the prompts. But we should expect participants' responses to be heavily weighted to System One responses, and thus to be irrelevant to the issue studied.[11]

6. Generalizing and How to Improve

Presumably, if some body of data is to be useful to help us adjudicate between different theories of consciousness, or between different theories of the concept of consciousness, the existence or non-existence (or the believed existence or non-existence of it) of a seemingly central thing like qualia should make a difference to that data. System One judgments about others' mental states will be the same whether or not qualia exist, or whether or not the folk believe that qualia exist. Data about these judgments is not terribly useful to deciding between theories of what consciousness is or what we conceptualize it as. Further, we should expect experimental studies involving judgments about others' mental states to largely gather elicit System One judgments. Thus, the problems I have raised here should generalize, and will be issues for a range of experimental research into consciousness via folk judgments.

To illustrate how this problem generalizes, let us briefly consider one other study. Joshua Knobe and Jesse Prinz's (2008) study on ascriptions of mental states to groups found that participants were willing to ascribe mental states such as believing, deciding, intending, knowing, or wanting to corporations. This is what we should expect from System One judgments; these mental states are not valenced or affective, and so should

---

[11] Since conducting the research that their original paper was based on, Sytsma and Machery have performed further tests partly intended to shed light on the cognitive processes behind ascriptions of mental states. See their "The Relevance of Folk Intuitions to Evaluating the Justification for the 'Hard Problem,'" from the Second Annual Online Consciousness Conference, http://consciousnessonline.wordpress.com/2010/02/19/the-irrelevance-of-folk-intuitions-to-the-%E2%80%98hard-problem%E2%80%99-of-consciousness/ They administered the "Cognitive Reflection Test" (CRT) to subjects who participated in a study similar to the one I criticize in my paper. The CRT is supposed to diagnose subjects' tendency to report System Two (conscious) judgments rather than System One (intuitive) ones; high CRT subjects are more likely to rely on System Two (Fredericks, 2005). Sytsma and Machery found no correlation between subjects' CRT scores and their judgments about mental states. They suggest that high CRT subjects were more likely to employ System Two judgments, and since their judgments did not differ significantly from those of low CRT subjects, Systems One and Two generate the same judgments about others' mental states.

This conclusion is too fast, however. A high CRT score is predictive of a willingness to engage System Two, but those who have high CRT scores do not engage System Two for all, or even most, of the judgments they make (otherwise they would be incapable of surviving). Subjects who score highly on the CRT test are those who use careful, conscious mathematical reasoning, rather than relying on intuitions about math (see Frederick 2005). However, we have generally been taught that math is difficult and counter-intuitive. Later research on the CRT test also indicates that making subjects experience difficulty when answering the questions – for example, putting them in a difficult to read font – dramatically raises scores (Alter, et al, 2007). So it seems that high CRT subjects are only more likely to engage System Two when they are presented with a task that they perceive to be difficult. Results from high CRT subjects in Sytsma and Machery's experiments only tell us about System Two judgments if these subjects perceived the task presented to them as difficult. If my argument about the perceived lack of difficulty of the task in Sytsma and Machery's study stands – if the task did not seem difficult to participants – then we should not expect even high CRT subjects to engage System Two, and the responses of high CRT subjects tells us nothing new.

be associated largely with behavior. It seems possible for corporations to behave as if they want, believe, intend, and so forth, and so System One should be willing to ascribe these mental states to groups in relatively ordinary circumstances. On the other hand, participants were not willing to say that corporations experienced joy, depression or pain. These are valenced states. Corporations do not (literally) look like things that should be able to experience these states – specifically, they lack the facial features that tend to trigger simulations of such states in participants' minds. System One should thus be unwilling to ascribe these states to corporations. Participants were also not very willing to say corporations are "experiencing a sudden urge," or "vividly imagining." Now, these are not clearly emotionally charged mental states, but at the same time it does not seem that corporations can behave in ways that we associate with imagining or having "sudden" urges – they cannot report to us that they are imagining, or look dreamily off into the distance, etc.[12] So these results are what we would expect from System One judgments.[13] Further, we should expect System One's to make these judgments whether or not qualia exist, and whether or not these states are truly qualitative.

Given this general problem, how should we gather data with which to study consciousness? One way to try to address the problem I have raised would be to ask participants to make judgments about their *own* mental states. If qualia exist, participants do have extensive experience of their own qualia, and it would make sense for System One to be able to use the presence or absence of qualia as a ground for ascribing mental states to one's self. There are two difficulties with this approach. First, it cannot be used to replicate studies that are specifically concerned with the mental states of things like groups, or cyborgs (e.g. Huebner, 2010). Second, it is very difficult to replicate the structure of other studies to gather judgments about our own mental states. Sytsma and Machery, for example, compare ascriptions of mental states to beings who are taken to be conscious and to those that are taken to not be conscious. This is not accidental, but is instead crucial to the argument they want to make. To replicate their study using data on participants' judgments about their own mental states, we would need to compare participants' judgments about their own conscious and unconscious mental states. However, it is in a certain sense hard to imagine ourselves being unconscious, and it is likely that participants make judgments about their unconscious selves just as they would make judgments about others. If this cannot be avoided, it would mean that System One judgments about our own mental states are just as likely to ignore qualia as System One judgments of others' mental states. Of course, the question of how we make judgments

---

[12] A similar point is made in Sytsma & Machery, 2009.

[13] Knobe and Prinz did find that subjects were willing to say that corporations *are* upset, rather than *felt* upset – at least, they reported that saying that corporations feel upset sounds more "weird" than saying they are upset. While this may seem surprising given everything I have said so far, this experiment is methodologically problematic. Apparently, subjects in it were given both sentences to compare: e.g. "Acme corp. is feeling upset" and "Acme corp is upset about the court's recent ruling." It's impossible to tell from this study if subjects were endorsing the claim that Acme corp. could be upset, or if they were simply reporting that this is less odd than talking explicitly about its feelings. Further, the clear contrast between the sentences is likely to generate bad data, as subjects will try to guess the answer they are supposed to give or that is expected by the experimenter, and more attention will be drawn the "feels" locution than might ordinary be paid to it. In the absence of better data, we should ignore this result. See also Sytsma & Machery, 2009, for further criticism of this specific study.

about our own, unconscious, mental states is an empirical one, and this needs further study.

Another way around this problem would be to study judgments from other mental systems, such as System Two. I am skeptical about this option. In order to do this, we need to give participants a reason to doubt their intuitive judgments, and to think they can judge better consciously. However, if we call participants' intuitions into question, this is likely to call their judgment about the issue more generally into question, given that it is natural to make these judgments using System One. This is likely to get participants to second-guess whatever answers they give, and generate bad data. Further, it does not seem to me that most people have thought much about how to categorize and ascribe mental states. This means that they are not likely to know of a better way to ascribe them other than just going with what intuitively seems right, or guessing.

We might look at ways to get the folk to articulate an awareness of their own qualia; if the folk cannot be brought to do so (as Sytsma and Machery suggest) then we have good evidence that there are no qualia. The danger here is that we will, in a sense, talk people into believing in qualia, rather than get them to talk about something they already believe in. After all, a great many philosophers believe in qualia; if there really are none, then there is something about thinking about philosophy of mind that tends to cause one to believe in qualia. Putting the folk in mind of the distinctions they need to talk about what we are interested in will always run the risk of taking from them that which makes them useful as participants.

7. Conclusion

Folk participants' judgments about the mental states of others are most likely to be generated by System One. System One will generate the same judgments whether or not we experience phenomenal consciousness. For this reason, studying either consciousness or our concept of consciousness through folk judgments about the mental states of others is largely inappropriate. I do not think that studying the hard problem via folk judgments is impossible, just that there are significant difficulties in this line of research, and current uses of these judgments is flawed. If we are to do experimental philosophy well, and I think we can and should, we must strive for a deeper understanding of our various mental systems and how they work. We must strive to tailor our experiments to employ the most appropriate mental systems for the task at hand, and to play to the strengths of these systems.

Acknowledgments

This paper could not be the paper it is without my discussions with Justin Sytsma and Bryce Huebner in the 2010 Online Consciousness Conference, and the invaluable help of Julia Staffel, Rob Rupert, and the students in my Experimental Philosophy seminar at C.U. Boulder.

References

Atler, A.L., Oppenheimer, D.M., Epley, N., & Eyre, R.N. (2007) Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology: General*, *136*, 569-576.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001) The 'reading the mind in the eyes' test revised version: a study with normal adults, and adults with Asperger Syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, *42*, 241-251.

Chalmers, D. (1995) Facing up to the problem of consciousness. *Journal of Consciousness Studies*, *2*, 200-219.

Chalmers, D. (1996) *Consciousness*. Oxford: Oxford University Press.

Frederick, S. (2005) Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*, 25-42.

Goldman, A.I. & Sripada, C.S. (2005) Simulationist models of face-based emotion recognition. *Cognition*, *94*, 193-213.

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

Kahneman, D. & Frederick, S. (2002) Representativeness revisited: Attribute substitution in intuitive judgment. In Gilovich, T., Griffin, D., Kahneman, D. (Eds). *Heuristics and Biases* (pp. 49-81). New York: Cambridge University Press.

Knobe, J. & Prinz, J. (2008) Intuitions about consciousness: experimental studies. *Phenomenology and the Cognitive Sciences, 7*, 67-85.

Lewicki, P., Hill, T., & Czyzewska, M. (1994) Nonconscious indirect inferences in encoding. *Journal of Experimental Psychology: General*, *123*, 257-263.

Lewicki, P., Hill, T., & Czyzewska, M. (1992) Nonconscious acquisition of information. *American Psychologist*, *47*, 796-801.

Sloman, S.A. (1996) The empirical case for two systems of reasoning, *Psychological Bulletin*, *119*, 3-22.

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

Sytsma, J. & Machery, E. (2010) Two conceptions of subjective experience, *Philosophical Studies,* 151, 299-327.

Sytsma, J. & Machery, E. (2009) How to study folk intuitions about phenomenal consciousness. *Philosophical Psychology*, 22, 21-35.