

# Philosophy of Science

## Problem Set #3

Ozoner Hansha

April 27, 2021

### Question 1

**Part a:** What's the hard problem of consciousness? What's the meta-problem of consciousness?

**Response:** The hard problem of consciousness refers to the so called explanatory gap between physical phenomena, like brain processes, and our phenomenal experience (e.g. qualia, mental states, etc.). The meta-problem of consciousness is the problem of why we think the hard problem is hard/real.

**Part b:** Could either problem conceivably be solved in a way that leaves the other unsolved? If so, sketch such a solution and say whether this shows that each problem is irrelevant to the other. If not, say why not.

**Response:** The meta-problem does not need to engage with concepts of qualia/phenomenal experience to be answered. It can be answered via purely physical means, presumably by considering the psychology of positing 'consciousness experience' in the first place. Whether this has a bearing on the hard problem depends. In an illusionist view, it might, as explaining why we have this illusion of phenomenal experience may suffice. However in other views, explaining why we *believe* we have phenomenal experience does not explain why we *have* this phenomenal experience, nor what it is.

**Part c:** Consider the dialogue between the realist and the illusionist on pp. 54-5 of Chalmers's paper on the meta-problem. Say whose reasoning you take to be more persuasive and why.

**Response:** It's hard to say which is more persuasive. While I might lean towards the illusionist side philosophically, both sides here seem reasonable in their stances. The dialogue comes off as if two brick walls are trying to push one another. That said, the realist's beliefs seem to be portrayed as more clearly 'dogmatic' than the illusionist's. The realist seems to simply assert that pain must be real, presumably since they observe it, despite conceding that observing pain is not evidence for it.

**Part d:** Consider the debunking argument on p. 47 of Chalmers's paper. Say whether you think this is a good argument. (You may draw on material from the paper in your response, though you need not.)

**Response:** I think it is a good argument. It basically makes the observation that all the phenomena we observe, including ourselves, are causally closed by physics. There is simply no room for consciousness to have any causal power over our actions and 'beliefs' (at least the non-qualia kind). If this is the case then, all our 'phenomenal intuitions,' as Chalmers puts it, just coincidentally line up with the physical part of ourselves. We are just philosophical zombies, but with qualia added on top. This makes our assertion that our qualia is real no different from a zombie's assertion that its qualia is real.

Of course the argument presented is a bit more terse, but assuming this is the idea it is appealing to, then I find it very convincing.

**Part e:** Do you think that argument is better poised to support strong illusionism or weak illusionism (see p. 49 for the distinction)? Say why or why not.

**Response:** While this argument seems to serve both forms of illusionism well, weak illusionism is probably more immediately served by it. This is because weak illusionism does not differ from realism as much as strong illusionism does, and this argument explains the gaps between realism and weak illusionism. Namely that consciousness doesn't have the ephemeral properties we normally ascribe to them and that these ascriptions are just a product of physical processes.

## Question 2

**Part a:** What is a meta-problem?

**Response:** A meta-problem is a problem about a problem. Of course, the meta-problem of consciousness is an example.

**Part b:** Formulate a problem and a meta-problem in another domain. Identify an analog of a solution to the meta-problem of consciousness for the meta-problem in that domain and say why the solutions differ in plausibility.

**Response:** Consider the following question posed by a hypothetical pre-Einsteinian physicist:

“Why does the precession of Mercury not obey the laws of physics?”

This is a reasonable question as the precession of Mercury is actually dependent on relativistic effects. And so, predictably, the meta-question we might ask here is:

“Are the ‘laws of physics’ this question refers to even correct?”

The original question is analogous to the hard problem as it asks why there is some explanatory gap between what we believe (our laws of physics/observations of a materialist universe) and what we see (Mercury’s precession/phenomenal experience). The meta-question is analogous to the meta-problem as it asks why we even believe there is a gap (are these laws even right?/why do we think there is phenomenal experience?).

In this case, the answer to the question is: no. General relativity is a (more) correct laws of physics and its acceptance resolves the first question. This is analogous to how illusionism as an answer of the meta-problem would clear up the hard problem. Of course, the validity of general relativity is much easier to demonstrate than illusionism.

**Part c:** Formulate a problem and meta-problem in a domain distinct from the one you chose for 2b. Explain how a conceivable solution to the meta-problem is relevant to solving the problem.

**Response:** The following question is still related to physics but more so in the domain of computer science & math, and I think it fits well so...:

“Are all physical phenomena computable (able to be simulated by a computer)?”

It would seem that our current understanding of physics would net the positive answer to this question. But objectors might point to the fact that our current understanding of physics is currently understood to be incomplete (gravity + quantum mechanics, dark matter, etc.). Indeed, they may object, there could be some unknown/unexplained phenomena that has some sort of uncomputable behavior (i.e. some behavior that can’t be predicted by a computer and thus can’t be predicted by humans).

Now consider a currently unaccepted theory like string theory. Such a theory perfectly explains the union of gravity and quantum mechanics as well as some other currently unexplained physical phenomena. It’s just not accepted, and rightly so, because of a lack of evidence. But the very fact that this is a plausible theory that exists and explains these phenomena in a coherent and calculateable (read: computable) way has bearing on the original question. Of course, string theory is no silver bullet, and leaves open other question to be sure, but just being able to come up with computable theories that explain phenomena seem to be enough to prove that said phenomena are computable. Whether or not the theories are ‘correct’.

I skipped over what the meta-question corresponding to this thought would be. I suppose its something like this:

“Are there even theories that show all physical phenomena to be computable?”

**Part d:** Describe a case in which the genetic fallacy is committed. (Don’t use examples from the link or trivial variations of them.)

**Response:** An example:

“The NYT printed an article, so everything in it must be true.”

Here the genetic fallacy is assuming that all printed media must be true by virtue of it being printed.

**Part e:** Explain why bringing a solution to a meta-problem to bear on a problem need not commit the genetic fallacy.

**Response:** It may seem that countering the hard problem via the meta-problem, e.g. questioning the hard problem on the basis of it being come up with by humans who evolved to believe in consciousness, may be committing the genetic fallacy. And indeed similar arguments could be made for meta-problems in general. However, this is not necessarily the case. The genetic fallacy is only committed if the consideration of a subject’s background is all that is considered in forming a judgement on the subject’s claim. Actually bringing a solution to the meta-problem would rely on evidence beyond just the questioning of humans’ belief in consciousness.

## Question 4

**Part a:** Schneider notes that some views of consciousness, if true, will make certain ways of “merging” AI with the mind impossible. Pick one view of consciousness and a possible method of AI “merging” that would be impossible if that view of consciousness were true. Explain. Now pick a view of consciousness and a possible method of AI “merging” that would be possible if that view were true. Explain.

**Response:** One theory of identity is bodily materialism. Here ‘you’ are just your body/brain, and as it changes over time it is still you. In this view merging with AI via say, copying your brain into a computer, would not work as that copy would not be ‘you’ as we have just defined. The real ‘you’ is that collection of matter particles the copy came from.

Another theory of identity is patternism. Here ‘you’ are the specific pattern of information that your physical body/brain implements. In this view, copying your brain to a computer would be just fine, as that ‘pattern’ is precisely what *you* are, the details of how it is implemented physically being irrelevant.

**Part b:** Explain the difference between qualitative identity and numerical identity. Use an example.

**Response:** There is a qualitative identity between two things if they share some properties. For example, a poodle and a chihuahua are both dogs and are thus qualitatively identical in that sense. Even further, two poodles are probably more qualitatively identical, in that same sense, than a poodle and a chihuahua.

Numerical identity is the complete and total qualitative identity of two things. This is a much more rigid notion of equality. For example,  $2 + 2 = 4$ . Real world examples are only the most trivial, for example an object, like an electron or person, at some point in spacetime is equivalent to itself at that same point in spacetime. Myself and myself two seconds from now are *not* numerically equivalent as there are differences between us (the particles in my body have evolved for 2 more seconds than my past self).

**Part c:** Numerical identity is transitive: if A is numerically identical to B and B is numerically identical to C, then A is numerically identical to C. When applied to persons, this means that a person at a certain time cannot be numerically identical to two different people at a different time. This seems like a problem for some versions of the psychological continuity view (including “Patternism”). Explain why.

**Response:** This poses a problem for any notions of personal identity that allow for duplication of the self, patternism being one key example. This is because, if all a person is is the information that comprises them and a digital copy of them could be made, then it could be arbitrarily copied. This would leave us with multiple copies of the same person but at different locations in space/time. These copies would not be numerically identical, just qualitatively identical. Where, then, is the ‘self’?

**Part d:** Which view of personal identity do you think is most plausible? Why?

**Response:** I think many of the problems materialist views of personal identity, like patternism, can be easily solved by stipulating that a person is the quantum information that comprises them. This is because quantum information cannot be destroyed and, crucially, cloned (no-cloning theorem). So, quantum teleporting oneself (like we have done with smaller collections of atoms) to another location, or even into a computer, doesn't pose the same star-trekkian problem as standard teleporters do because there is no instant in time where the same information is in two places at once. Indeed such a case would violate physics.

That said, while QM may be a silver bullet for materialist views of personal identity, I don't buy personal identity regardless. It would seem to me that notions of a 'self' are byproducts of our evolution, particularly 'our' psychology. 'Our' of course referring to the masses of flesh being puppeted by the statistical fluctuations of equally fleshy brains.

**Part e:** Does your view of personal identity make any potential mechanisms for AI "merging" impossible? Explain.

**Response:** I think not taking seriously the notion of a self would make merging with an AI, say by copying yourself into a computer, a non-issue? Indeed people will always implicitly see themselves as a singular self, and so non-self-ists will still look after their well-being somewhat, and probably still take issue with a reduplication situation. But that can easily be avoided by stipulating you quantum teleport yourself into the computer rather than copy yourself in. At least, I'd have no problem in assuring myself that I'm the same person who teleported into the computer.