

Title : Self-evidencing conscious experience and vicious circularity

Abstract (150 words) :

The meta-problem of consciousness aims to explain the particularity of our intuitions about consciousness and how they trigger conceptual issues such as the hard problem of consciousness. I propose in this article that these stem from a basic function of the brain : self-evidencing explanation. To make sense of its sensory inputs, the brain is believed to build and test models of the state of the world based on sensory information (Hohwy, 2016). This self-evidencing process has been proposed to describe the type of inference performed by consciousness (Friston, 2018). I will show how this situation is viciously circular and prevents from proving the existence of conscious experience or explaining it without presupposing its existence. I will show how it accounts for the particularity of our intuitions about consciousness and thus propose an solution to the meta-problem of consciousness using a formally defined process at the core of conscious inference.

1. Explaining conscious experience through conscious self-evidencing

1.a. Consciousness is self-evidencing

The hard-problem of consciousness is explaining the relationship between conscious experience and physical processes. This task has been deemed hard because of the strength of the intuition that physical processes cannot possibly explain conscious experience, also called the explanatory gap (Levine, 1983). This problem has been opposed to easy problems, ie explaining features of consciousness that do not conflict with our intuitions about consciousness. The meta-problem of consciousness is explaining the particularity of our intuitions regarding conscious experience and has been considered an easy problem (Chalmers, 2018). It holds a particular place among easy problems since it could help to understand the root of what makes the hard-problem of consciousness hard. To explain our intuitions about conscious experience, I start with examining the type of inference performed by the brain.

The inference process performed by the brain has been described as being self-evidencing explanation (Hohwy, 2016). Self-evidencing occurs

whenever “the information or assumption that the explanandum is true provides an indispensable part of the only available evidential support for one of the explanans statements” (Hempel, 1965:373). For example, from the presence of smoke and the knowledge that smoke is caused by fire, the hypothesis that a fire is occurring can be inferred. In this case, the evidence that the smoke is present or the assumption of the presence of smoke is necessary to infer the presence of the fire.

To infer a hypothesis, the brain relies on the construction of models of the states of the world based on sensory information. The brain relies both on sensory data to build models and to test them against sensory evidence. Evidence are used to confirm or reject a hypothesis. These models have been proposed to support conscious processing when used to actively infer through building and testing hypotheses about the states of the world causing sensations (Hobson and Friston, 2016; Hobson 2018).

1.b. Self-evidencing conscious experience is problematic

Explaining conscious experience requires thus to infer the state of the world causing conscious experience and to be able to gather evidence that can allow to confirm or reject this model. In this case, conscious experience is the explanandum of the self-evidencing process. If conscious processing is self-evidencing, the hypothesis that conscious experience exists or evidence of conscious experience are necessary as part of the explanans to explain conscious experience.

One can accept the hypothesis that conscious experience exists but refuse to use evidence about conscious experience to explain conscious experience. This corresponds to explaining conscious experience using information that do not presuppose consciousness, only in terms of physical processes. Nevertheless, explaining conscious experience uses self-evidencing about conscious experience and as such, it uses evidence that is gathered leading to conscious experience.

Another one can deny the hypothesis that conscious experience exists and accept to use evidence produced by the self-evidencing. This situation triggers a particular epistemological situation as recently pointed by Francois Kammerer with the “illusion meta-problem” (Kammerer, 2016). Provided the hypothesis that conscious experience is an illusion, ie does not exist, testing

the absence of conscious experience would require having the conscious experience of not having a conscious experience, which is paradoxical.

The “illusion meta-problem” introduced here is close to the meta-problem of consciousness because it aims to explain the strength of our intuition coming against the refutation of the existence of conscious experience. I will now demonstrate how self-evidencing precisely accounts for the particularity of such situation and show in a third part its relevance for our intuitions about consciousness.

2. The vicious circularity of self-evidencing explanation

2.a. The circularity of self-evidencing

Self-evidencing has been criticized to be a case of circular inference process since the information about the occurrence of the explanandum is used as an essential part of its explanans. For example, the presence of smoke is used both to infer the hypothesis of the presence of a fire and as evidence for its presence.

The circularity can be limited through testing hypotheses based on independent sensory evidence, as noted by Hempel, “an acceptable self-evidencing explanation benefits, as it were, by the wisdom of hindsight derived from the information that the explanandum event [i.e. the smoke] has occurred, but does not misuse that information so as to produce a circular explanation” (Hempel 1965:373). For example, the reliability of the hypothesis of a fire happening can be increased or decreased through collecting independent evidence, e.g. through checking if oxygen levels are decreasing in the surrounding atmosphere.

In this case, the epistemic circle is benign (Lipton 2004). The circle turns vicious, however, under a specific condition, namely when doubts arise about the occurrence of the evidence itself. Thus if someone doubts the occurrence of the smoke, it would be misguided to defend it by appeal to the fire, given there is only evidence for the hypothesis that there is a fire, if it really was smoke in the first place.

2.b. The specificity of explaining conscious experience as viciously circular

The hypothesis that the smoke is caused by the fire does not suppose a self-evidencing process that will infer and test its existence. In this case, independent evidence can be easily gathered through self-evidencing to test the presence of smoke. But for conscious experience, since it results from conscious processing which is self-evidencing, acquiring evidence about the existence of conscious experience requires the use of self-evidencing and thus lead to conscious experience. Thus, the existence of conscious experience cannot be independently tested, ie without using information about the occurrence of conscious experience.

Even if the hypothesis that the evidence exists cannot be tested, can such evidence be explained using self-evidencing ? To avoid circularity and be able to be tested independently, no evidence using self-evidencing can be gathered since no information about the occurrence of evidence must be used. Thus, self-evidencing requires to accept the assumption that the evidence exists as part of one of the explanans statement. Therefore, evidence cannot be explained only by hypotheses that do not presuppose its existence. As such, conscious experience cannot be explained with hypotheses that do not suppose the existence of conscious experience, ie only by physical explanations.

One could still deny the hypothesis that the evidence exists and accept to use information about the presence of evidence. Nevertheless, the hypothesis that evidence does not exist is in contradiction with the use of evidence using self-evidencing. Thus, the hypothesis that the evidence does not exist cannot be tested by the self-evidencing process without contradiction between its existence and its use. This is the case of illusionists. By contrast, if one accepts the hypothesis that evidence exists, such hypothesis cannot be independently tested but no contradiction arises. Thus the hypothesis remain that conscious existence does not exist, and illusionists might be true, but testing such hypothesis is in contradiction its use during self-evidencing.

3. Self-evidencing consciousness as a solution to the meta-problem of consciousness

3.a. A topic-neutral account of the meta-problem of consciousness

My main argument is thus that :

P1. consciousness is self-evidencing

P2. a self-evidencing process cannot prove the existence of its evidence and explain its evidence by the use only of hypotheses that do not suppose its existence

C. Thus, any conscious being cannot test the existence of its conscious experience or explain it without supposing its existence.

I would like to highlight in this last section how this formalization accounts for our intuitions about conscious experience and makes several non trivial predictions on the role they play in the debate about the existence of conscious experience. Self-evidencing explanation sheds light first on why the meta-problem of consciousness is “a problem of explaining phenomenal reports in topic-neutral terms”, i.e. that do not involve conscious experience (Chalmers, 2019). Indeed, to avoid the problem of circularity arising from self-evidencing, any acceptable explanation would require to be inferred from independent evidence, i.e. topic-neutral terms, than the phenomena which it tries to account for, i.e. phenomenological notions. As such, self-evidencing is formulated in topic-neutral terms (Hohwy, 2016; Friston, 2018).

3.b. Epistemic and metaphysical intuitions at the core of the hard-problem

I showed how explaining the evidence of self-evidencing using self-evidencing represents a particular epistemic situation for which self-evidencing is viciously circular. Any hypothesis to explain conscious experience must not presuppose the existence of the conscious experience to be able to be tested in a way that avoid circularity. Nevertheless, I showed that conscious experience cannot be explained with only with hypotheses that do not suppose its existence. This particular situation makes sense of our epistemic intuitions about the explanatory gap between conscious experience and physical notions that do not suppose the existence of conscious experience.

Self-evidencing offers indeed an account of the illusion meta-problem through explaining how the fact of denying the existence of conscious experience could not even be intuitively conceived because it would contradiction with the necessity to presuppose the existence of conscious experience to form an hypothesis. The strength of the intuition of the existence of conscious experience could thus be explained by the fact that the

hypothesis that conscious experience exists does not come with such contradiction in its formation. Because conscious experience cannot be reduced to physical notions that do not suppose conscious experience and can be explained without contradiction only if it presupposes its existence, the metaphysical intuition that conscious experience exists appears to be fundamental. Nevertheless, self-evidencing predicts that the hypothesis that conscious experience exists cannot be tested. Thus the question whether conscious experience exists remains unsettled according to self-evidencing and eliminativists and illusionists might be right.

Self-evidencing explains that the epistemic and metaphysical intuitions come from the fact that, to explain conscious experience, the same self-evidencing process is used to generate hypotheses and to test them. The first-person perspective allows to form and gather evidence about the existence of its own conscious experience but this situation triggers a vicious circular inference. As an alternative, the approach to treat conscious phenomena in a third-person perspective has been proposed by Dennett introducing heterophenomenology (Dennett, 1993). This strategy exemplifies how to avoid circularity because phenomenal reports are not produced by the same self-evidencing process than the one used to explain them. Thus, phenomenal reports can be explained as sensory data and tested in the same way as physical hypotheses. Such approach solves methodological problems in decoupling our intuitions from the way to test them. But it does not prevent intuitions about conscious experience from emerging based on the use of self-evidencing in a first-person perspective, nor to solve the problem of providing independent evidence about the existence of conscious experience using self-evidencing.

3.c. The knowledge and modal argument and the third-person approach

In order to try to provide support for the existence of conscious experience from a third-person perspective, two types of arguments have been proposed : the knowledge argument and the zombie-thought experiment. I will show how self-evidencing can account for these intuitions and for their role in proving the existence of conscious experience. While the knowledge argument addresses whether one can find independent evidence that conscious experience exists and is not reducible to physical explanations, the second one

addresses whether the hypothesis that conscious experience does not exist can be formed without contradiction.

The problem of providing independent evidence that conscious experience exists and is not reducible to physical processes has been discussed with Mary's room thought experiment (Jackson, 1982). Mary is a scientist in possession of all third-person knowledge regarding a physical state of the world but has never experienced it, e.g. the red color. The argument is that her first sensory encounter with the red color would still provide her with new information that she did not possess previously that would be conscious experience. Self-evidencing predicts that when Mary is seeing a red color for the first time, independent evidence resulting from self-evidencing is gathered to test her models of the red color. Such inference results in the conscious experience of red. Hence, the intuition of additive evidence linked to the existence of the conscious experience of red is explained by self-evidencing as inferring the cause of a newly encountered sensory situation. Importantly, the evidence obtained in the first-person perspective is here discussed in topic-neutral terms as knowledge does not suppose the existence of conscious experience. The knowledge argument plays an important role in the debate by the possibility to provide independent evidence that conscious experience exists without presupposing the existence of conscious experience.

The problem whether the hypothesis that conscious experience does not exist can be formed without contradiction has been addressed through the zombie thought experiment, i.e. the intuition that there can be a functionally equivalent counterpart of a conscious being but devoid of conscious experience (Chalmers, 1996). Self-evidencing predicts effectively that I can conceive a zombie in a third-person perspective, since no contradiction exists between the hypothesis that the conscious experience of the zombie does not exist and my conscious experience that is required when using self-evidencing. The third-person perspective allows thus to formulate without contradiction the hypothesis that conscious experience does not exist, but does not address whether such hypothesis is actually true and can be tested. The self-evidencing framework informs thus both our understanding of the modal intuition about conscious experience and its role played in inferring the existence of conscious experience.

4. Conclusion

Self-evidencing is an inference process that has been proposed to be at the core of the conscious process. I highlighted a particular epistemological case for which self-evidencing forms a vicious circular inference that cannot be solved through collecting independent evidence. I showed how this explains our intuitions about the irreducibility of conscious experience to physical states of the world that do not presuppose conscious experience. It also accounts for the fact that the existence of conscious experience cannot be proven or disproven from a subjective point of view as well as the strength of the intuition that conscious experience is fundamental. It gives an insight on the role that these intuitions play in the debate about conscious experience and the constraint of explaining conscious experience in topic-neutral terms to avoid the circularity of self-evidencing. Self-evidencing explanation gives thus an account for the particularity of our intuitions about conscious experience and their importance in the philosophical debate about the hard-problem of consciousness. Relying on consciousness as self-evidencing, I show how this account offers a solution to the meta-problem of consciousness and its explanatory power as a framework for consciousness (Hohwy, 2016; Friston, 2018).

5. References :

- Chalmers, D. (2018). The meta-problem of consciousness. *Journal of Consciousness Studies*, 25(9-10), 6-61.
- Dennett, D. C. (1993). *Consciousness explained*. Penguin uk.
- Friston, K. (2018). Am I Self-Conscious?(Or Does Self-Organization Entail Self-Consciousness?). *Frontiers in psychology*, 9.
- Hobson, J. A., Friston, K. J. (2014). Consciousness, dreams, and inference: the cartesian theatre revisited. *Journal of Consciousness Studies*, 21(1-2), 6-32.
- Hohwy, J. (2016). The self-evidencing brain. *Noûs*, 50(2), 259-285.
- Jackson, F. (1982). Epiphenomenal qualia. *The Philosophical Quarterly* (1950-), 32(127), 127-136.
- Kammerer, F. (2016). The Hardest Aspect of the Illusion Problem--and How to Solve it. *Journal of Consciousness Studies*, 23(11-12), 124-139.
- Nagel, T. (1974). What is it like to be a bat?. *The philosophical review*, 83(4), 435-450.