

## Peer Review

# Review of: "No Consciousness? No Meaning (and no AGI)!"

Francesco Abbate<sup>1</sup>

1. Sapienza University of Rome, Italy

The paper presents important ideas on the grounding problem, in most cases clearly and according to relevant references. While I share many of them, I believe that it does not demonstrate its main thesis nor does it really argue it. My suggestion is to reduce its scope, since if the central thesis is removed, the paper still offers several valid insights. For example, the statement according to which “the paper argues that while adding sensory inputs and embodying AI through sensorimotor integration with the environment might enhance its ability to connect symbols to real-world meaning, this alone would not close the gap between syntax and semantics” may be valid and informative for reasons that go beyond the central thesis, as well as for reasons that are easier to explain. Alternatively, the central thesis must be explained and discussed much better, but also, and most importantly, demonstrated. Sentences like “The key difference between even the most advanced AI and a person with sensory impairments is that the absence of sight in humans is compensated for by other subjective experiences, like sound, touch, taste, and smell. These experiences enable the indirect semantic comprehension of visual concepts even if direct visual perception is lacking, as the representation is supported by other forms of conscious experience—something a machine cannot replicate” do not really help to explain such a thesis. First of all, because it reduces the subjective experience to the senses, which is simplistic. Rather, it should be said that blind human beings retain the ability to perceive the world through the other senses and through these perceptions (which are something more than simple sensation) have a subjective experience of the world. Secondly, because it remains unclear how we pass from perceptions to conscious states (and to which ones) and why machines are not capable of having them (they certainly aren’t, starting from the fact that they are not capable of perceiving the world). Also, the paper is littered with repetitive, apodictic statements that do not demonstrate the hypothesis, such as, for example: “Phenomenal consciousness and its whole psychological dimension cannot be sidestepped, as it is inextricably intertwined with meaning-making” (p.7); “because understanding is an aspect of phenomenal consciousness itself” (p.18);

“To understand something implies being conscious of it” (p. 19); “True understanding is inherently rooted in and inseparable from consciousness” (p.19); “meaning and conscious experience are inseparable” (p.23).

I will first discuss the problems related to the central thesis, as they are more critical, and then, in order of importance, the other critical points.

The central thesis is, as the title suggests, that it is not possible to achieve artificial general intelligence (AGI) without first achieving consciousness. Apart from the fact of completely conceiving AGI on human intellectual capacities without ever questioning whether we might have forms of intelligence other than human intelligence that can do without human supervision, this thesis presents both fundamental problems, i.e., relating to what the author means by consciousness, and methodological problems, since the distinction between hypotheses and findings is weak, if not completely absent.

Firstly, the author never defines what he means by consciousness. This is not, in fact, an on-off concept but a broad spectrum along which it is possible to situate various higher cognitive functions of the animal kingdom. To address the thesis that he proposes necessarily means addressing the question of what levels of consciousness are necessary to be able to understand human language. Furthermore, the author often confuses the question of understanding human language with that of the ability to attribute meaning tout court, concluding that “meaning and conscious experience are inseparable” (p.23). This is a direct consequence of the lack of definition of what is meant by consciousness. In fact, it is not true that a consciousness of the human type is necessary to be able to attribute meanings to things, and most likely, conscious states of any other type are not necessary either.

Let's take the concept of proto-self (Damasio). It is an unconscious level deriving from the internal visceral states of the body. Already from this level, it is possible to derive adaptive capacities to environmental stimuli that are fundamental for animals, the definition of their own goals, and therefore the definition of their own meanings (it is, by the way, something very similar to the biosemantic perspective to which the author refers and that can be applied even to plants, which it is hard to consider conscious). Reptiles, for example, to which we can attribute a proto-self, have attack and defense systems against external objects that are intrinsically significant to them – as prey or as threats. Furthermore, we can identify even more complex levels of self that are still unconscious. Panksepp talks about a core-self and identifies it as a self capable of experiencing primary emotions through which the individual can have an adaptive emotional experience and derive an unconscious memory from it (Le Doux).

Furthermore, the author never demonstrates that full human consciousness is necessary for machines to understand human language. I am doubtful that this evidence can exist, but I will gladly change my mind if the author presents any arguments to support his thesis, which he does not at the moment. His argument is based on the fact that, since it is true for human beings, then it must also be so for machines. However, first, it would be necessary to demonstrate why it must necessarily be so for machines as well and why it cannot be otherwise, at least in part. For example, machines are capable of performing tasks that humans are not able to perform in the absence of conscious states, such as carrying out complex mathematical operations that require the use of working memory (on the contrary, we know that human beings are able to carry out basic tasks or simple mathematical operations such as  $2+2$  without needing to be fully conscious). As for human beings and their linguistic abilities, it is correct that these depend on their conscious states, but it is also correct to hypothesize that the opposite is true. In fact, more complex and significant conscious states may have developed thanks to and through the use of language, that is, through the development of what Damasio calls the autobiographical self that goes hand in hand with the development of language capacities. In short, we should succeed in the arduous task of explaining how we pass from the more basic states of consciousness, such as those shown by chimpanzees, who can be hypothesized to be capable of being self-aware of themselves (Gallup; Lewis), to more complex states that involve a capacity for symbolic self-recognition. This shift toward a symbolizing self involves the ability to be aware of one's own mental states, such as thoughts, emotions, and feelings, and in this process, more complex symbol manipulation abilities may play a crucial role. In this sense, therefore, it wouldn't be exclusively the basic levels of consciousness that enable an understanding of human language but other cognitive abilities, possessed by humans and not by chimpanzees (although one could argue that the determining factor may have been quantitative rather than qualitative, i.e., that the transition from apes to humans occurred thanks to the selection of more intelligent individuals who managed to find different adaptation strategies to the environment compared to those of other primates).

The paper then engages also with the concept of subjective experience. I think it's better to address the issue of language understanding using this concept rather than that of consciousness. It is certainly correct to assume that it is not possible to understand language or determine its meanings in its absence and that this therefore represents a prerequisite for those. However, the author never really explains what is meant by subjective experience, nor does he offer any insight into how this is determined. It would be very useful if he used some references in this regard, referring to constructs that are rooted in the

cognitive sciences and not exclusively in Chalmers' philosophy of mind (limiting the concept of subjective experience to saying that it concerns the "what it is like" is, in my opinion, a bit simplistic, and it doesn't really help us in understanding the gap between the human experience and the lack of experience of artificial beings). The suggestions I gave regarding the different types of self may be useful in this regard.

Furthermore, some concepts are introduced, such as that of embodiment, which are left without references and are not explained. What does the author mean by embodiment? The simple fact of having a physicality, and in this sense, even a laptop would have it, or something else? In this second case, the sentence 'Embodiment may be a necessary condition, but it is not sufficient to transform unconscious comprehension into conscious comprehension' is vague and adds not much, especially if included in the conclusion.

Finally, I suggest shortening the conclusion also by eliminating references to a whole series of theories such as enactivism, extended mind theory, the free energy principle, etc. As it is, this list is of little use, and its presence in the conclusion is out of place.

Some editorial considerations: Please insert numbers in front of the titles of each section. In the introduction, it is necessary to introduce the sections briefly.

## **Declarations**

**Potential competing interests:** No potential competing interests to declare.