Does a computer experience placebo?

Over 70 years ago, Alan Turing asked himself, and the world, a very simple question: "Can machines think?"[1] Such a proposition has sparked endless debate about the nature of computation, and how analogous are machines to human beings. Never has this conversation been more prevalent than right now (12/23) with the radical ascension of Artificial Intelligence (AI). It seems as if the personability of these machines plagues great doubt into both layfolk and experts.[2] Right now, there are countless studies about how ChatGPT writes like a human,[3] reasons like a human,[4] and even has similar moral values to humans.[5] These results have proved to have no shortage of intellectual camps that argue either for or against AI's similarity to humans. In fact, there are so many of these arguments that they engendered various theories to support their conclusions. On the for side, you have computer & cognitive scientists that argue that mental states – and therefore cognition – is computational.[6] Those against, argue the converse; that mental states are not computational, and thus, computers/computational **alone** cannot explain cognition in its entirety.[7]

This paper is not designed to explore the depths and validity behind both camps and I will therefore only summarize their strongest points. On the for side there is an abundance of evidence that the brain can be replicated computationally. Many labs have been able to replicate human perception using neural networks, which are computational models that can learn any category that can be thrown at them.[8] Additionally, the field of computational neuroscience has demonstrated that the brain itself can be modelled computationally.[9] These findings provide a solid logical and empirical basis that both the structure and function of the brain operate computationally. Conversely, these models have their limits. The most notable limit is that they *are* computational theories. Hence, they are limited in the axioms in which they are based on. This is most clear in the now 30 year old problem posited by Stevan Harnad: The Symbol Grounding Problem (SGP). In sum, the SGP states that all computers do is manipulate symbols (computation), and they do not have any meaning derived from their symbol maniuplations, as symbol manipulation is the only thing computation can do. Thus, in order to have "meaning", or otherwise known as *subjective experience, sentience,* and *feeling*, one **must** have a body. Simple computational mechanisms will not engender feeling.[10] This argument is often misunderstood,[11] as meaning is grounded in sensory input. Computers, and machines alike, do not have sensory

---

[1] https://web-archive.southampton.ac.uk/cogprints.org/499/1/turing.html
[2] Citation needed about chatgpt
[3] Citation needed
[4] Citation needed
[5] Citation needed
[6] https://users.cs.northwestern.edu/~ian/What%20is%20computation.pdf
[7] Citation needed.
[8] Citation needed.
[9] https://plato.stanford.edu/entries/computational-mind/?ref=jack-chong#NeuNet
[10] https://eprints.soton.ac.uk/250382/1/symgro.pdf

[11] Cite podcast episode with harnad

input like biological systems do, they only have computational inputs and outputs. Hence, they are neither necessary nor sufficient to explain consciousness/feeling.[12]

Despite this clear limitation, the current debate still goes on as to whether computers are sentient, most notably AI. This debate is most clear in the work around ChatGPT, and other Large-Language Models (LLMs). Just last year, a Google engineer claimed that their AI was "sentient",[13] and there is no shortage of scientific work being done on ChatGPT across the globe.[14] Even if the SGP proves that there is a clear gap between humans (autonomous sensorimotor beings) and computers (machines), the conversation and *understanding* is still muddied. This ultimately means that the arguments, both for and against, are not strong enough, as neither side is being convinced from the other; ultimately leading to research that self-segregates in its theory versus integrating. As such, I propose a new argument against robot-sentience: that because computers cannot experience the placebo (or nocebo) effect, they are not sentient.

Why the placebo effect; what is so special about it? In this case, the placebo effect (or the nocebo effect, its counterpart) is a good example to prove by contradiction that computers are not sentient because it is a richly studied subject. It is testable, measureable, and replicable in humans and animals,[15] it is mapped out neurochemically,[16] and it is both a generic and specific psychological state.[17] It demonstrates a highly *unique* and *subjective* and yet specific biological event. Yet, it cannot be programmed into a computer. A computer cannot experience a placebo effect (nor a nocebo effect). Indeed, the mere existence of the placebo effect cannot serve as a measure for sentience, but it is something that sentient beings *experience*, and therefore, it can be lumped into the category of "parts that make up sentience".

This is an invariant distinction between machines and humans. As mentioned earlier, much of the discussion around ChatGPT is whether it "understands" language. The word "understand" is what muddies the water because the against camp would argue that understanding must have the *feeling* of understanding to understand. Whereas the for camp would argue that ChatGPT understands because it uses language correctly. Given the mutual irreducibility of these ontologies, it is clear that new argument needs to be proposed to further the discussion. This is why I propose that because a computer cannot experience the placebo effect, it cannot be sentient.

---

[12] Cite church/turing thesis.
[13] https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/
[14] Cite Harnad talks & papers. The bulk of this research is being done with whether ChatGPT "understands" the language that it is trained on. But this is foolhardy.
[15] Undergrad paper
[16] Undergrad paper
[17] Ctitation needed most likely