Review of: "Do Androids Dread an Electric Sting?"

Yew-Kwang Ng[1]

1 Monash University

Potential competing interests: No potential competing interests to declare.

Review [for *Qeios*] of the paper:

Do Androids Dread an Electric Sting?

By Izak Tait and Neşet Tan

Reviewer: Yew-Kwang Ng, Emeritus Professor of Economics, Monash University, Australia;

Part-time Professor, School of Economics, Peking University, China.

Kwang.ng@monash.edu

This paper addresses an interesting and potentially (but not yet) relevant and important question. It argues that

- 'we would need to create a framework for protecting artificial entities' (abstract) when (not 'if') AI we create will be capable of having pain and even 'mental suffering' (p. 12).
- Moreover, this will happen in a matter of just 'in fifty years' time or only five years' (Abstract).
- Further, they will have 'the potential to experience more pain and suffering than any other living entity on Earth' (Abstract).

On these issues involving the hard problem of consciousness, no one can say anything certain, as no scholars on earth (including scientists and philosophers) have any clue (not even by 0.01%) on how material brains can generate subjective consciousness, even for ourselves. So I cannot dismiss the three bulleted points above are certainly false. However, in terms of the balance of probability, I would say that these three points are wide overstatements, if not sensational.

For the first point, if we change the 'when' into 'if', I do agree with the need for protection. However, I think the timing of this is not 50 or 5 years; rather, it will likely be 5 centuries or even never. Many readers may regard this as being too pessimistic, given the speed in the progress of computer and AI technologies. Let us consider this issue.

Whether AI may become conscious depends on how subjective consciousness is generated. If consciousness is just a function of the purely materialistic brain. AI consciousness must be possible, at least in principle. If we build something similar to our brains, it should possess consciousness as we do. However, even so, I think the time taken to instantiate AI consciousness will still take centuries, not just decades to realize. The computers and robots we build are not based on simulating our brains which are too complicated to simulate. As mentioned above, we know almost nothing

on how our own consciousness is generated by our own brains. How to make androids capable of having consciousness is thus much difficult than the authors of the paper under review and likely most people in the AI cycles believe.

Secondly, the simple materialist position mentioned above may not be true; rather, dualism may be correct. If consciousness is not just a function of the purely materialistic brain, but also involves some essential mental element, androids, being purely materialistic, may never be conscious. Furthermore, the studies in parapsychology (defined broadly to include studies of mediumship, memories of past lives, near-death experiences, etc.), especially over the last few decades, provides many compelling evidences suggestive of postmortem survival (Ng 2023). This suggests that we may survive our biological death; our bodies may die, but our souls may survive. If so, this suggests that AI consciousness may be well nigh impossible. Nevertheless, I agree with the view (p. 12) 'better to be safe than sorry', so that if sufficiently strong evidences are available suggestive of AI consciousness, we should not be too demanding in asking for conclusive evidence before taking actions in providing protection of AI welfare, just as it should be in the case of animal welfare protection. Better to over-protect than under-protect. However, the actual situation is the opposite. We should move in the direction of more protection of animal welfare, and much could be done at negligible or even negative costs (considering over consumption of meat for health purposes) on humans (Ng 2016).

Given that androids do have consciousness and is capable of enjoyment and suffering (i.e. being affective sentients), then the reason why do we have to protect them as discussed in Section 3 of the paper under review, is clearly the moral badness of causing unnecessary suffering, as pain, suffering, unhappiness are intrinsically bad, and enjoyment and happiness are intrinsically good (Ng 2022, ch. 5).

Some specific points may also be discussed. Bottom of p. 4, 'some evidence showing that fish feel only the former [reactions to negative stimuli via nociception] and not the latter [subjective qualitative experience of pain]'. This is controversial; see a discussion in *Animal Sentience* 2016 on the view of Brian Key (cited in the paper under review) :

https://www.wellbeingintlstudiesrepository.org/animsent/vol1/iss3/1/ .

Bottom of p. 6, 'No researcher has yet found conclusive evidence that any AI model available today has phenomenal consciousness'. There is not only no conclusive evidence, there is not even suggestive evidence.

On p. 7, nine (latter focussing on five) attributes (building blocks) of consciousness are identified. I find that some of these attributes are too demanding and some are inadequate for affective feelings. e.g. if something perceives pain but has no memory of previous pain, it is still affective sentient. On the other hand, a computer may be programmed to 'Generating novel information via inferences of incoming information', but that is unlikely to entail consciousness.

On p. 8, 'We can already say thatcurrent AI models have met the criteria for two of the five Building Blocks: Perception and Data-Output' (underline added). In my view, there are only data-output, there is no perception, at least not in the subjective sense relevant for sentience.

References

NG, Yew-Kwang (2016). How welfare biology and common sense may help to reduce animal suffering,*Animal Sentience,*

1.7: 1. Target article for peer commentary.

http://animalstudiesrepository.org/animsent/vol1/iss7/1/

NG, Yew-Kwang 2022 . *Happiness: Concept, Measurement, and Promotion*, Springer, open access at: https://link.springer.com/book/10.1007/978-981-33-4972-8

Ng, Yew-Kwang (2023). Could artificial intelligence have consciousness? Some perspectives from neurology and parapsychology. *AI & Society*, *38*(1), 425-436.