Review of: "Questioning the Moratorium on Synthetic Phenomenology"

Trevor Thwaites

Potential competing interests: No potential competing interests to declare.

Thank you for the opportunity to read your paper. I found your argument sound, and the language comfortable to read, rather like a personal discussion. You treat Metzinger with respect even as you dismiss some of his claims. What follows are thoughts and literary encounters I have had since first accessing your paper, additionally stimulated by the Dreyfus brothers' notion of **cognitive simulation** in their work *Mind Over Machine* (1986, p.67), in which a project collected various ways humans solved problems and fed these into the machine – not unlike an early Large Language Model. Moving forward forty years, we now have excursions into genomics where biology as digital information and digital information as biology produce biodigitalism in nanoscience. The National Science Foundation (NSF) is also reporting on research where gene sequencing technologies are paired with artificial intelligence.

The contemporary emergence of technologies is taking us into the post-digital era and the fields of*biodigital convergence* and quantum theories. This is perhaps why a moratorium to prevent AI machines from suffering may already be too late, for the quantum world is not binary. The next leap forward in computing is likely to see a convergence of advances in co-design modelling and simulation, data analytics, machine learning, and artificial intelligence.

I note that scoping papers from Policy Horizons Canada (2022) refer to the complete integration of biological and digital entities in which new hybrid forms of life are created through the integration of biological components in digital technologies (Peters, 2023, *Education, Philosophy and Theory*, p. 1628).

Stepping back to the 80s, Jacques Ellul in*The Technological System* (1980, Continuum, pps 335 – 8) discusses the autonomy of technology in regard to values and morals. Note that technologies, for Ellul, included war machines encountered in his own experiences during WWII. The following five points are Ellul's, but I have abridged what is several pages in the book.

Analyze five aspects:

1. Technology does not progress in terms of a moral ideal, it does not seek to realize values, it does not aim at a virtue or a good.
2. Technology does not endure any moral judgement. Morality judges moral problems. Technology should judge/question itself because non-technological factors get in the way.
3. Since technology does not support any ethical judgement, it does not tolerate being halted for a moral reason, "situational morality" is quite convenient for putting up with anything. However, the autonomy of technology is so well-

assured that now technology is turning into a judge of morality.

4. Modern man takes for granted that anything scientific is legitimate and, in consequence, anything technological. To challenge becomes suspect, and it is technology that now validates scientific research.
5. Being self-justified, technology normally becomes justifying. The ethics is built on concrete givens, for it is primarily an experienced ethics of the behaviour required for the technological system to function well, giving it a vast superiority over the other moralities of being truly experienced.

Questions I pose for Metzinger's dilemma are:

Can an AI platform become a substitute for a sensing, living body? What perceptions can it make of the spatial world?

What capacities does AI need to develop and refine data, and how might the physical machine have an effect on the psychic? What phenomenon of causality enables the machine to 'suffer,' and how might the concept of suffering differ between machines and humans?

If we can turn human know-how into computable AI knowing-that, what happens when the facts or circumstances change?

Finally, and not wishing to sound trite, I ask: Is a hypochondriac 'suffering' machine a possibility?