

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

Michail Ploumis<sup>1</sup>

<sup>1</sup> Hellenic Military Academy

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

'The article embarks on a journey into the realm of AI consciousness and its potential implications for welfare and legal protection. It discusses the concept of "Building Blocks of consciousness," suggesting that certain capabilities could lead to an AI experiencing feelings with a phenomenal character. While this is an intriguing proposition, the discussion could have been significantly enriched by providing more concrete examples or case studies to illustrate these concepts. Furthermore, a more detailed exploration of the philosophical and scientific underpinnings of these "Building Blocks" would have added much-needed depth to this section.

Transitioning towards potential milestones that AI might need to reach to qualify for welfare status, the authors highlight a troubling paradox - that AI could potentially feel pain without having the agency to do anything about it. This is a critical point that deserves a more thorough exploration. The authors could have delved deeper into this paradox, exploring its implications for AI development and ethics. Furthermore, a broader range of viewpoints or theories on this matter could have been included to provide a more balanced perspective.

However, the section of the article that explores the reasons why we should consider the welfare of AI is particularly disappointing. The authors raise some thought-provoking questions about the potential ethical implications of creating conscious AI. However, their exploration of these questions is superficial. This section feels like a missed opportunity, as the authors could have provided a more comprehensive discussion, drawing on existing literature or ethical frameworks. A more nuanced understanding of this complex issue could have been achieved by exploring different ethical theories, discussing potential scenarios, or even considering the potential societal and legal implications of recognizing AI welfare.

The article then discusses the cognitive attributes that future AI would need to be at the same level as animals. The authors identify Perception and Data-Output as two of the five Building Blocks that current AI models have met, and point to Inference and Meta-Representation as the key milestones. While this provides a roadmap for future AI development, the authors could have further elaborated on the implications of these milestones. A more detailed discussion on how these milestones could be achieved, the potential challenges that might arise, or the ethical dilemmas that could be faced as AI continues to evolve would have added more depth to this section.

In the final section of the article, the authors attempt to draw conclusions from their discussion. However, this section feels somewhat rushed and lacks the depth and detail present in the rest of the article. The authors could have provided a more comprehensive summary of their discussion, highlighting the key points and their implications. They could have also provided a more detailed discussion of potential future research directions, providing a roadmap for further exploration of

this complex and important topic.

In conclusion, while the article tackles a complex and important topic, there is significant room for improvement in terms of depth and breadth of discussion. The authors have laid a foundation for discussions about AI consciousness and welfare, but the exploration of the topic could have been more comprehensive and nuanced. The article serves as a starting point, but further research and more detailed analysis would be beneficial in this rapidly evolving field.'

We recommend: Michail Ploumis, "AI Weapon Systems in Future War Operations; Strategy, Operations and Tactics," *Comparative Strategy* (41, no. 1, 2022),