

ROBOTS AND HUMANS: FROM THE PERSPECTIVE OF NEURO-ETHICS.

Tibor Szolcsányi

University of Pécs, Medical School, Department of Behavioural Sciences

■ ABSTRACT

Advances in artificial intelligence, particularly in deep learning systems, have renewed debates about the possible moral and legal status of intelligent machines. These discussions extend beyond the question of future AI rights and directly engage considerations about the neuroethical foundations of human dignity and human rights. Because contemporary AI architectures are, in certain respects, functionally analogous to aspects of human brain organization, it becomes necessary to clarify which features of neural functioning are ethically relevant for the attribution of moral status.

This essay examines one influential theory used to justify human and animal rights: Tom Regan's account of inherent moral worth. It analyzes Regan's notion of being an experiencing subject of a life and offers a phenomenological interpretation that emphasizes the role of phenomenal consciousness in human motivation and preference formation. The paper argues that, although intelligent machines may exhibit complex goal-directed behavior, there is currently insufficient justification for attributing moral rights to them. At the same time, the analysis demonstrates how debates about AI rights deepen our understanding of the neuroethical foundations of human dignity.

Keywords: neuroethics, artificial intelligence, Tom Regan, phenomenal consciousness

<https://doi.org/10.63154/CETR2025.1-6>

Introduction

The functioning of the human brain has ethical significance as evidenced by the widespread medical and legal acceptance of the whole-brain death definition of death (Bernat, 2005). This definition implies that biological personhood is determined—either fully or at least in part—by the brain’s capacity to coordinate bodily functions. This raises a fundamental question: is it possible to identify the specific attributes of the human brain that make each person intrinsically worthy? In other words, can it be rational to ground

the concept of human dignity and rights in the functioning of the nervous system?

The complexity of this question necessitates consideration of another related issue, which may offer some preliminary insights. The internal operations of the most advanced AI-based robots exhibit certain similarities to the functioning of natural nervous systems. This raises the critical question: under what conditions should a brain-inspired intelligent machine be considered as an entity deserving rights?

I turn now to this second question.

Critical Anthropomorphism: a conceptual and methodological tool

The question of whether intelligent machines and robots should be considered as entities with rights is a complex ethical and philosophical issue. As artificial intelligence and robotics evolve, the possibility of granting rights to machines must be examined within a structured theoretical framework. To find an appropriate theoretical framework, it can be useful first to introduce the concept of critical anthropomorphism, which refers to the cautious attribution of human-like traits to robots based on their observed behaviors. Naturally, a certain degree of anthropomorphism is unavoidable when examining the functioning and behavior of AI (for review, see Li & Suh, 2022), as AI is specifically designed

to perform human-like cognitive tasks, such as symbol-processing, memory, perception, among others. Moreover, AI capable of machine learning employs multilevel data-processing mechanisms that, in some relevant respects, resemble natural neural networks. As Arleen Salles (2020) and her colleagues emphasize, neuroscience has clearly inspired AI-research, but the reverse is also true because AI-research has inspired neuroscientists to better understand how the human brain works. It is therefore unsurprising that not only lay persons tend to use to kind of naive AI-anthropomorphism, but AI experts also tend to use anthropocentric language when analyzing the behavior and function-

ing of intelligent machines. (Salles, A., Evers, K., & Farisco, M. (2020)). However, unlike naive anthropomorphism, which assumes that any human-like behavior in machines implies consciousness, critical anthropomorphism serves as a methodological tool for analyzing robotic agency and cognition.

The concept of critical anthropomorphism originates in debates on animal welfare issues. In that context, critical anthropomorphism is the view that the compassionate treatment of animals must be grounded in objective knowledge—utilizing scientific methods—regarding the evolution, behavior and internal processes of the animals and species in question (Donnelley & Nolan, 1990; Morton, Burghardt & Smit, 1990). Naturally, debates exist in the literature concerning the role and significance of critical

anthropomorphism (see e.g., Karlsson, 2012). However, from a philosophical standpoint it is much clearer that critical anthropomorphism is essential for formulating strong arguments against certain versions of panpsychism, such as a view that suggests all physical objects can experience pain.

If a robot merely simulates human emotions and decision-making through advanced programming but lacks any true internal experience, it remains a sophisticated tool rather than a rights-bearing entity. However, if a brain-inspired robot's behavior indicates a deeper cognitive structure—one that allows for learning, adaptation, and self-initiated action—then denying its moral status might be unjustified. Ultimately, the answer depends on the theoretical framework adopted.

Human rights and the philosophical basis for moral consideration

To determine whether an intelligent machine or a robot should be granted rights, it is useful to draw upon established theories regarding human and animal rights. Tom Regan's concept of inherent moral worth is particularly relevant in this context. (Regan, (2004). Regan argues that an entity deserves rights if it is an experiencing subject-of-a-life—an individual capable of experiencing its environment, recognizing the effects of its own behavior, and pursuing personal interests based on future-orient-

ed desires. In Regan's view, therefore, we do not have sufficient reason to attribute inherent moral value to an entity merely because it is sentient; other attributes play a decisive role in this determination. Consequently, if a robot can demonstrate characteristics suggesting that the robot is an experiencing subject-of-a-life, such as the ability to form goals, avoid harm, or express preferences, then denying its moral status could be inconsistent with Regan's theoretical framework.

The key question, however, is whether a robot's actions are the result of genuine internal preferences or merely predefined responses to stimuli. To examine this question more thoroughly, I will offer an interpreta-

tion of Regan's view on what it means to be an experiencing subject-of-a-life. An analysis of phenomenal consciousness and its role in human experiences and motivations will be central to my interpretation.

The phenomenal aspect of experiences and human preference-formation

Phenomenal consciousness is commonly defined as the aspect of consciousness associated with the concept of qualia. Undoubtedly, human experiences often possess qualitative features that define how sensations, emotions, and moods feel to an individual—or as Thomas Nagel (1974) phrased it, “what it is like” for a person to have conscious experience. Even within the same sensory modality, experiences can differ significantly in their qualitative character—for example, variations in colors, tastes, odors, or musical sounds.

A defining characteristic of phenomenal consciousness is its subjectivity. Only the individual undergoing a particular experience has privileged, first-person access to its qualitative content. In contrast, external observers can only have indirect, inference-based access to the phenomenal properties of another's experience (Nagel, 1974). This asymmetrical access, and the resulting individual uniqueness of phenomenal states, is perhaps the most important attribute of phenomenal consciousness (Haladjian & Montemayor, 2016)).

David Chalmers (1997) famously introduced the “Hard Problem of Consciousness” to emphasize a fundamental explanatory gap in contemporary neuroscience and philosophy of mind: while science has made substantial progress in mapping the neural correlates of conscious experiences, it remains unclear how and why specific patterns of neural activity give rise to phenomenal states. Nevertheless, human motivations are often directed toward phenomenal states rather than merely toward states of affairs. Consider, for instance, one of the most fundamental human motivations: the desire for safety. While appropriate external conditions are typically necessary for people to experience a sense of safety, these circumstances often serve merely as a means to the subjective experience of what it feels like to be safe—or to ensure the subjective feeling of safety for loved ones. The phenomenal aspect of experiencing safety, therefore, is in many cases an inherent component of the genuine human motivation for safety.

Numerous other examples can illustrate the

relevance of phenomenal states in achieving, maintaining, and ensuring preferred individual experiences. According to my best knowledge, the first author who clearly argued for the functional role of phenomenal consciousness in the human cognitive system was Uriah Kriegel (2004). In a recent article, Cleeremans & Tallon-Baudry, (2022) also emphasized the role of phenomenal states in human decision-making and preference formation. By

contrast, although intelligent machines can exhibit goal-directed behavior, there is a consensus among AI researchers that, at the current stage of AI development, there are no sufficient practical reasons to assume that the behavior of intelligent machines results from subjective experiences or phenomenal consciousness (see, e.g., Haladjian, H. H., & Montemayor, C. (2016); Ng, G. W., & Leung, W. C. (2020); Woodward, P. (2022).

Concluding remarks

If we accept my phenomenological interpretation of what it means to be an experiencing subject-of-a-life, then there is currently insufficient justification for granting rights to intelligent machines and robots within the framework of Tom Regan's theory. From a neuroethical perspective, it is also crucial to recognize that the neural processes oc-

curing in the human brain can give rise to—and correlate with—subjective experiences and the phenomenal aspects of human consciousness. It may be possible to develop a conception of human dignity that reflects the fact that, under appropriate neural conditions, human beings are experiencing subjects-of-a-life.

Acknowledgement

This work was funded by the Medical School of the University of Pécs.

Some parts of this essay—specifically, the final paragraph of Section 1 and all paragraphs in Section 2—were written with the assistance of AI, based on my own PowerPoint slides.

The prompt used was: "Create a coherent academic text based on the uploaded slides." In addition, various parts of the essay were refined with AI support to enhance American English grammar and academic style, in part following AI-assisted editorial feedback.

Bibliography

- Bernat, J. L. (2005). *The concept and practice of brain death*. *Progress in brain research*, 150, 369-379. [https://doi.org/10.1016/S0079-6123\(05\)50026-8](https://doi.org/10.1016/S0079-6123(05)50026-8)
- Chalmers, D. J. (1997). *The conscious mind: In search of a fundamental theory*. Oxford Paperbacks.
- Cleeremans, A., & Tallon-Baudry, C. (2022). *Consciousness matters: phenomenal experience has functional value*. *Neuroscience of consciousness*, 2022(1), niac007. <https://doi.org/10.1093/nc/niac007>
- Donnelley, S., & Nolan, K. (1990). *Special supplement: animals, science, and ethics*. *The Hastings Center Report*, 20(3), 1-32.
- Haladjian, H. H., & Montemayor, C. (2016). *Artificial consciousness and the consciousness-attention dissociation*. *Consciousness and Cognition*, 45, 210-225.
- Hildt, E. (2019). *Artificial intelligence: does consciousness matter?*. *Frontiers in psychology*, 10, 1535. | <https://doi.org/10.3389/fpsyg.2019.01535>
- Karlsson, F. (2012). *Critical anthropomorphism and animal ethics*. *Journal of Agricultural and Environmental Ethics*, 25(5), 707-720. DOI:10.1007/s10806-011-9349-8
- Kriegel, U. (2004). *The functional role of consciousness: A phenomenological approach*. *Phenomenology and the Cognitive Sciences*, 3(2), 171-193.
- Kriegel, U. (2019). *The value of consciousness*. *Analysis*, 79(3), 503-520. <https://doi.org/10.1093/analys/anz045>
- Li, M., & Suh, A. (2022). *Anthropomorphism in AI-enabled technology: A literature review*. *Electronic Markets*, 32(4), 2245-2275. <https://doi.org/10.1007/s12525-022-00591-7>
- Morton, D. B., Burghardt, G. M., & Smith, J. A. (1990). *Critical anthropomorphism, animal suffering, and the ecological context*. *The Hastings Center Report*, 20(3), S13-S13.
- Nagel, T. (1974). *The Philosophical Review*. *What is it Like to Be a Bat*, 435-450. <http://www.jstor.org/stable/2183914>
- Ng, G. W., & Leung, W. C. (2020). *Strong artificial intelligence and consciousness*. *Journal of Artificial Intelligence and Consciousness*, 7(01), 63-72. <https://doi.org/10.1142/S2705078520300042>
- Regan, T. (2004). *The case for animal rights*. Univ of California Press.
- Salles, A., Evers, K., & Farisco, M. (2020). *Anthropomorphism in AI*. *AJOB neuroscience*, 11(2), 88-95. <https://doi.org/10.1080/21507740.2020.1740350>
- Woodward, P. (2022). *Consciousness and rationality: The lesson from artificial intelligence*. *Journal of Consciousness Studies*, 29(5-6), 150-175. DOI: <https://doi.org/10.53765/20512201.29.5.150>