



Existenz

Volume 17, No 1, Spring 2022

ISSN 1932-1066

Artificial Systems – Agents or Processes?

Catrin Misselhorn

University of Goettingen, Germany
catrin.misselhorn@uni-goettingen.de

Abstract: In this brief replik to Jörg Noller's essay "Robots, Emotions, and Interobjectivity" that is included in this issue of *Existenz*, I highlight three central concepts in an effort to contrast some differences in addressing human agents' actions versus artificial agents' actions. These concepts include processes, life-worlds, and virtual realities. While I agree with many of Noller's observations regarding the first two books of my trilogy, our positions significantly differ regarding the role of humanism in matters concerning artificial intelligence.

Keywords: *Künstliche Intelligenz und Empathie*; artificial intelligence; empathy; relational artifacts; life-world; virtual reality; subject-object distinction.

I am grateful to Jörg Noller for his insightful discussion of some of the issues that my two first books raise.¹ I am disposed to agree with many observations he is making, yet I want to sharpen some of his points so that they become a genuine alternative to the view that I am defending in my books, being aware that this comes at the cost of sometimes overstressing the differences between our views. I want to explicate these differences regarding three central concepts: (1) agents *versus* processes, (2) life-world, and (3) virtual reality.

Regarding (1), it seems to me that his position is based on an ontology that reduces everything to impersonal processes. Only given such a strong interpretation of processes can his view become a genuine alternative to mine and there can exist such a seamless transition between humans and artificial systems as Noller conceives of it. In contrast, in my view the distinction between agents and artifacts is

deeply rooted in humans' conceptual scheme and this is why one also thinks of artificial systems in these terms. One categorizes them as agents on account of their degree of autonomy and intelligence that distinguishes them from mere artifacts.

Moreover, humans considering themselves as agents is fundamental to the way in which they understand themselves and this understanding is crucial for ethics. Issues of responsibility, for instance, only arise with respect to agents not with respect to impersonal processes. Noller's attempt "to avoid dualistic subject-object divisions" (*REI* 20) hence comes at a high theoretical and ethical cost for it means, to speak in Sellarsian terms, a fundamental revision of humans' manifest image and their place in it. Whereas I believe that the subject-object schema needs to be extended by intermediate categories such as artificial agents and relational artifacts to the end of fitting properly into artificial systems while leaving intact the categories of subject and object at the end of the spectrum, Noller wants to get rid of the distinction wholesale.

¹ Jörg Noller, "Robots, Emotions, and Interobjectivity," *Existenz* 17/1 (2022), 19-25. [Henceforth cited as *REI*]

I suggest that his project ultimately amounts to a revisionist metaphysics that would alter humans' conceptual scheme profoundly. This is problematic with respect to (2) the role of the life-world. Noller aims at a "life-worldly interpretation of artificial systems" (*REI* 19). Yet, his emphasis of the life-world is not compatible with the revisionist character of his process metaphysics. Perhaps it is correct that the world consists at its most fundamental metaphysical level of processes. At the manifest level of the life-world, however, the difference between agents and artifacts remains nevertheless central. Although arguably humans do effortlessly interact with artificial systems, it remains an interaction between two distinct agents and does not amount to a fusion of humans and artificial systems as Noller is suggesting it (at least not in my vigorous interpretation of his view).

A logical continuation of his project would include (3) virtual reality. Humans' conceptual scheme relies on their perception of and interaction

with physical reality as it appears in the manifest image. And this physical reality is characterized by boundaries between different objects as well as objects and agents. This fact resists the seamless transition between humans and artificial systems that Noller's view requires in respect to physical reality. It could only be overcome by virtualizing reality, since in virtual reality there are no physical objects and the boundaries between virtual objects are only as strict as they are programmed to be. The culmination of this ideal of eliminating the physical boundaries between humans and artificial systems would be to use not just virtual-reality-glasses or other types of non-invasive technologies to enter virtual realities but computer-brain interfaces or even the uploading of the brain to guarantee what Noller calls "a holistic integration of intelligence performances" (*REI* 23). If this is right, then Noller's view might turn into a post-humanist vision in the end, or to put it the other way round: Humanism presupposes a strict distinction between subject and object.