



HoloFoldit and Hologrammatically Extended Cognition

Cody Turner¹

Received: 10 October 2022 / Accepted: 23 October 2022
© The Author(s) 2022

Abstract

How does the integration of mixed reality devices into our cognitive practices impact the mind from a metaphysical and epistemological perspective? In his innovative and interdisciplinary article, “Minds in the Metaverse: Extended Cognition Meets Mixed Reality” (2022), Paul Smart addresses this underexplored question, arguing that the use of a hypothetical application of the Microsoft HoloLens called “the HoloFoldit” represents a technologically high-grade form of extended cognition from the perspective of neo-mechanical philosophy. This short commentary aims to (1) carve up the conceptual landscape of possible objections to Smart’s argument and (2) elaborate on the possibility of hologrammatically extended cognition, which is supposed to be one of the features of the HoloFoldit case that distinguishes it from more primitive forms of cognitive extension. In tackling (1), I do not mean to suggest that Smart does not consider or have sufficient answers to these objections. In addressing (2), the goal is not to argue for or against the possibility of hologrammatically extended cognition but to reveal some issues in the metaphysics of virtual reality upon which this possibility hinges. I construct an argument in favor of hologrammatically extended cognition based on the veracity of virtual realism (Chalmers, *Disputatio* 9:309–352, 2017) and an argument against it based on the veracity of virtual fictionalism (McDonnell and Wildman, *Disputatio* 11:371–397, 2019).

Keywords Extended cognition · Mixed reality · Microsoft HoloLens · Metaphysics of virtual reality

How does the increasing integration of mixed reality devices into our cognitive practices impact the mind from a metaphysical and epistemological perspective? In his innovative and interdisciplinary article, “Minds in the Metaverse: Extended Cognition Meets Mixed Reality,” Paul Smart addresses this underexplored question, arguing that the use of mixed reality devices (specifically, the Microsoft HoloLens)

✉ Cody Turner
cturne22@nd.edu

¹ Technology Ethics Center, University of Notre Dame, 216 O’ Shaughnessy Hall, Notre Dame, IN 46556, USA

represents a technologically high-grade form of extended cognizing. In particular, Smart demonstrates how a hypothetical application of the HoloLens, which he calls the HoloFoldit case, qualifies as an instance of extended cognition from the perspective of neo-mechanical philosophy. There are numerous conceptual payoffs of this intellectual endeavor, which is sure to be of great interest to both philosophers and computer scientists.

This commentary aims to (1) carve up the conceptual landscape of possible objections to Smart's argument and (2) elaborate on the possibility of *hologrammatically extended cognition*, which is supposed to be one of the features of the HoloFoldit case that distinguishes it from more primitive forms of cognitive extension. In tackling (1), I do not mean to suggest that Smart does not consider or have sufficient answers to these objections. In addressing (2), the goal is not to argue for or against the possibility of hologrammatically extended cognition but to reveal some issues in the metaphysics of virtual reality upon which this possibility hinges. I construct an argument in favor of hologrammatically extended cognition based on the veracity of virtual realism (Chalmers, 2017) and an argument against it based on the veracity of virtual fictionalism (McDonnell & Wildman, 2019). Rather than a criticism of Smart's argument, then, this commentary functions more as a short companion piece.

The Microsoft HoloLens is a mixed reality headset that enables users to project three-dimensional virtual objects onto their perceptual fields such that the virtual objects appear to be fixed in physical space. The HoloFoldit case is a hypothetical application of the Microsoft HoloLens in which users can actively manipulate virtual objects to solve the protein folding problem, which concerns the question of how a protein's amino acid sequence relates to its three-dimensional atomic structure. The process involved in solving the protein folding problem is what Smart labels the *Protein Structure Prediction (PSP) process*. While hypothetical, the HoloFoldit case is based upon an existing system called "Foldit" that enables users to engage in the PSP process. Instead of having the three-dimensional protein molecule represented on a two-dimensional computer screen, the HoloFoldit renders the molecule as a hologram in physical space. And instead of having to use a keyboard to participate in the PSP process, the HoloFoldit allows users to partake in the process via a combination of hand gestures, eye gazes, and voice commands.

After presenting the HoloFoldit application, Smart proceeds to contend that the case qualifies as a form of extended cognizing because it satisfies three key mechanistic criteria for extended cognition related to (1) *the problem of cognitive status* (is the relevant mechanism a cognitive mechanism?), (2) *the problem of constitutive relevance* (does the mechanism constitutively contain both a human component and an extra-organismic component?), and (3) *the problem of cognitive ownership* (can the mechanism be attributed to the human component?).

The first possible objection that one might levy against Smart's argument is fairly straightforward.

Objection 1 (O1): Deny that the HoloFoldit case is an example of extended cognition from the perspective of neo-mechanical philosophy.

It is, of course, possible to pursue O1 on the grounds that the extended mind thesis is false and should be replaced by something like the embedded mind thesis (e.g., Rupert, 2004), but put this possibility aside. The more compelling defense of O1 comes from the proponent of the extended mind who thinks that the HoloFoldit case fails to satisfy one or more of the mechanistic criteria for extended cognition. First, said proponent might aver that the HoloFoldit case fails to satisfy the cognitive status criterion because the PSP process is not a bona fide cognitive process. According to Smart, the question of whether the PSP process qualifies as cognitive hinges on whether the PSP process satisfies what Clark and Chalmers (1998) call the parity principle (i.e., whether we would consider the PSP process to be a cognitive process if it were performed inside the head of a single human individual). Smart takes it to be obvious (or at least highly intuitive) that the PSP process satisfies the parity principle. I am inclined to agree with him, but an objector could, in principle, push back on this front, especially considering that questions surrounding “the mark of the mental” are notoriously controversial.

Alternatively, one might defend O1 by proclaiming that it is impossible to resolve the problem of constitutive relevance for the HoloFoldit application because the case is hypothetical in nature. As Smart explains, so-called interventionist approaches to the problem of constitutive relevance maintain that “issues of constitutive relevance are resolved by experimental interventions that reveal the relationship between component-level and phenomenon-level variables” (Smart, 2022: p. 17). Since the HoloFoldit is not an existing technological system, the device cannot be subject to such experimental intervention. Consequently, the problem of constitutive relevance arguably cannot be solved with respect to the device. In response to this possible objection, Smart says that while interventionist approaches to the problem of constitutive relevance make sense in the context of scientific investigation, they are not applicable from an engineering perspective because the role of the engineer, unlike the scientist, is *not* to discover mechanisms via empirical experimentation but to create them. Thus, any engineer tasked with writing code for the HoloFoldit app would already possess knowledge of the mechanistic underpinnings of the system, meaning that nothing would be gained epistemically from subjecting the system to empirical scrutiny.

I find these responses by Smart to be persuasive, but it is worth noting that he could have circumvented the above two objections entirely had he focused on a mixed reality application other than the HoloFoldit. For example, Logg et al. (2017) have developed an application for the Microsoft HoloLens that they call *HoloFEM*, which allows users to “define and solve a physical problem governed by Poisson’s equation with the surrounding real-world geometry as input data” (Logg et al., 2017: 1). Unlike the HoloFoldit and the PSP process, the HoloFEM clearly satisfies the parity principle (and by extension the cognitive status criterion) because it pertains to mathematical processes of calculation that are sometimes (or are at least capable of being) neurally realized by human biological brains. Moreover, the HoloFEM is an existing (as opposed to hypothetical) mixed reality application and so is immune to the possible interventionist objection discussed above. The upshot is that while the HoloFoldit case is novel and connects to contemporary interest in protein

folding problems, it is arguably not the best example for the purposes of illustrating that mixed reality devices can facilitate extended cognition, as there are more clear-cut examples of mixed reality-based extended cognition that avoid some of the objections faced by the HoloFoldit case.

Even if it is granted that the HoloFoldit case is a form of extended cognizing, one might object to the idea that it is a type of *human-centered* extended cognition.

Objection 2 (O2): Claim that the HoloFoldit case is an instance of machine-centered extended cognition, not human-centered extended cognition.

O2 relates to the problem of cognitive ownership. Essentially, one might contend that in the context of the HoloFoldit and the PSP process, the locus of agential control and responsibility (and therefore ownership) is properly attributed not to the human operating the HoloLens device but to the device itself. This notion of machine-centered extended cognition (or “human-extended machine cognition”) derives from Paul Smart (2018), who avers that some kinds of human–machine interaction involve human minds being integrated into a machine (AI)-based cognitive apparatus instead of machines being integrated into a human-based cognitive apparatus. Smart briefly considers O2 in footnote 14 (p. 23) but dismisses the idea that the HoloLens owns the PSP process, given that the human user is primarily responsible for dictating the trajectory of the process when operating the HoloFoldit app. He remarks that the HoloFoldit app does not “recommend that the user undertake specific actions in response to the current problem state. If this were to be the case—if, for example, the HoloFoldit app, were to start guiding the human user through the problem-solving process by recommending specific actions—then I suspect our ownership-related intuitions might begin to shift” (p. 23).

This point is intuitive enough, but to provide a more effective response to O2, we need a more thorough, methodical account of what cognitive ownership entails, one that allows us to more easily determine when an extended cognitive mechanism is machine-centered versus human-centered. For example, many contemporary smartphone apps such as TikTok constantly recommend specific actions and engage in algorithmic nudging as a means to keep users plugged into the platform. In such cases, especially when social media addiction is involved, there is a genuine sense that the machine intelligence is controlling the human user and not the other way around. Does this mean that some existing cases of human-smartphone interaction qualify as instances of machine-centered extended cognition? More conceptual work needs to be conducted on the problem of cognitive ownership to offer a sufficient answer to this question.

Finally, even if the objector concedes that the HoloFoldit case is a bona fide example of human-centered extended cognition, they might seek to undermine the theoretical value and purported novelty of the argument.

Objection 3 (O3): Claim that the HoloFoldit case is not sufficiently distinctive compared to oft discussed low-tech cases of extended cognition.

Smart anticipates O3 by highlighting two distinctive features of the HoloFoldit case: (1) the idea that HoloFoldit represents an example of *internet-extended cognition* because it involves integrating online elements into human cognitive processes and (2) the idea that HoloFoldit represents an example of *hologrammatically extended cognition* because it involves integrating holographic (or virtual) elements into human cognitive processes. While the notion of internet-extended cognition has already been subjected to considerable conceptual analysis (e.g., Smart, 2017; Smart & Clowes, 2021), the notion of hologrammatically extended cognition has thus far been largely neglected in the extended cognition literature. Since this is supposed to be one of the features of the HoloFoldit case that sets it apart from low-tech cases of extended cognition, it is worth briefly elaborating on the concept of hologrammatically extended cognition.

The possibility of hologrammatically extended cognition essentially hinges upon whether virtual objects possess causal power. The following argument supports this proposition.

P1. In order for an extra-organismic entity X (e.g., a hologram) to be a part of a (human-centered) extended cognitive mechanism Y (e.g., the protein folding process), X must be constitutively relevant to Y.

P2. An entity X is constitutively relevant to a mechanism Y only if X is a real entity that possesses genuine causal power.

C1. Therefore, in order for an extra-organismic entity X (e.g., a hologram) to be a part of a (human-centered) extended cognitive mechanism Y (e.g., the protein folding process), X must possess genuine causal power.

P1 is just an expression of the constitutive relevance criterion for extended cognition, whereas P2 is a claim about what constitutive relevance entails. The veracity of P2 can be gleaned from Smart's discussion of the concept of "causal betweenness": "constitutive relevance is to be understood as a form of "causal betweenness." What this means is that a component must form part of a causal path that connects two events that together delimit the temporal bounds of the explanandum phenomenon (i.e., the events that mark the beginning and ending of the explanandum phenomenon)" (p. 17: 609–615). Thus, the holograms (or virtual objects) involved in the HoloFoldit case are constitutively relevant to the explanandum phenomenon (i.e., the PSP process) only if the holograms possess causal power.

Whether virtual objects possess causal power is a question about the metaphysical status of virtual objects. There are two main theories concerning the metaphysical status of virtual objects: *virtual realism* and *virtual fictionalism*. Virtual fictionalism maintains that virtual objects are fictional objects that do not really exist (McDonnell & Wildman, 2019) whereas virtual realism holds that virtual objects are real, mind-independent entities with robust causal powers (Chalmers, 2017, 2022). Depending upon one's perspective on this debate, the above argument can be lengthened into either an argument for or against hologrammatically extended cognition. The virtual realist version of the argument in favor of hologrammatically extended cognition can be run as follows.

The Virtual Realism Argument for Hologrammatically Extended Cognition

- P3. Virtual objects possess genuine causal power if virtual realism is true.
- P4. Virtual realism is true.
- C2. Therefore, virtual objects possess genuine causal power.
- C3. Therefore, hologrammatically extended cognition is possible.

There is a strong *prima facie* case for virtual realism and the possibility of hologrammatically extended cognition. As Smart discusses on pages 25–26, virtual objects of perception arguably are better candidates for extended cognition than physical objects of perception. Smart observes that Palermos (2014) rules out the possibility of physical objects of perception partly constituting an extended cognitive process on the grounds that such objects do not participate in reciprocal causation (which Palermos takes to be a necessary condition for extended cognition). Palermos avers that the act of walking around a tree while looking at it, for instance, does not qualify as a type of extended cognizing because the perceptual interaction only features a one-way causal relationship: while the tree causally affects the perceiver, the perceiver does not in turn causally affect the tree. Unlike physical objects of perception, such as trees, virtual objects of perception appear to satisfy Palermos' reciprocal causation condition. Mixed reality devices must constantly track the user's head position to generate virtual objects that feature perceptual constancy. This means that the virtual objects of perception in the HoloFoldit case are causally sensitive to the user's movements because they must be continuously updated by the device based on the user's physical location.

There are further nuances though. Whether hologrammatically extended cognition is possible may depend not just on if virtual realism is true but also on what version of virtual realism is correct. Chalmers (2017) defends a version of virtual realism that he calls *virtual digitalism*, according to which virtual objects are digital objects, meaning that they are to be identified with corresponding data structures in the requisite computing device. Chalmers presents two closely related arguments for virtual digitalism (the argument from causal powers and the argument from perception), both of which hinge on the idea that virtual objects have genuine causal power that they derive from the digital objects that ground them. In footnote 16 (p. 25), Smart rejects virtual digitalism, claiming instead that virtual objects are best conceptualized as photonic objects (i.e., objects of light), not digital objects. We might call this view, which is plausibly a version of virtual realism, *virtual photonism*.¹

¹ Smart astutely points out in private correspondence that it is debatable whether virtual photonism is best construed as a type of virtual realism or virtual fictionalism. This is because photons exist in the realm of quantum mechanics and there are competing interpretations of quantum mechanics. Smart says that the Copenhagen interpretation provides *prima facie* support for the idea that virtual photonism is a variant of virtual fictionalism, whereas Pilot wave theory supports the idea that the view is a variant of virtual realism. Smart also suggests that reflection upon the notion of virtual photonism serves to challenge the idea that there is a neat binary distinction between virtual realism and virtual fictionalism in the first place.

It is beyond the scope of this commentary to adjudicate which version of virtual realism is more friendly to the possibility of hologrammatically extended cognition. There are many considerations and hidden complexities here that demand further philosophical investigation. For instance, one might ask: does hologrammatically extended cognition require that virtual objects (or holograms) are able to casually affect physical objects *as well as other virtual objects*? Put differently, to satisfy the reciprocal causation condition for extended cognition, do the virtual objects involved in the HoloFoldit system need to be able to participate in reciprocal causation with both conventional physical objects and other virtual objects that make up the system? If so, then virtual photonism is plausibly incompatible with hologrammatically extended cognition since photons cannot casually affect one another (but instead pass through each other without being affected). By contrast, virtual digitalism is consistent with the idea that there exists reciprocal causation between virtual objects assuming digital objects can casually interact with one another.

There are also arguments against virtual digitalism, however. McDonnell and Wildman (2019) distinguish between two forms of virtual digitalism: *strong virtual digitalism* (according to which virtual objects are identical to digital objects) and *weak virtual digitalism* (according to which virtual objects are dependent on, but distinct from, digital objects). They reject the strong version of the view because it is susceptible to what they call “the cross-play problem” and aver that the weak version cannot accommodate genuine causation between virtual objects: “the weak option is too weak: virtual causation is either a pseudo-process, unfit to sustain the causal commitments of the picture proposed, or is excluded in favor of its digital base” (McDonnell & Wildman, 2019: p. 889). After arguing against virtual digitalism, McDonnell and Wildman construct a positive argument in favor of virtual fictionalism inspired by Kendall Walton’s theory of fictionality (Walton, 1990). Anyone sympathetic to this view (which they call “virtual walt fictionalism”) might pose the following lengthened version of the above argument against hologrammatically extended cognition.

The Virtual Fictionalism Argument Against Hologrammatically Extended Cognition

P3*: Virtual objects do not possess genuine causal power if virtual fictionalism is true.

P4*: Virtual fictionalism is true.

C2*: Therefore, virtual objects do not possess genuine causal power.

C3*: Therefore, hologrammatically extended cognition is impossible.

This argument has an intuitive pull, but it is unclear whether the veracity of virtual fictionalism is incompatible with the possibility of hologrammatically extended cognition. For instance, McDonnell and Wildman’s argument in favor of virtual fictionalism is focused solely on causal interactions between virtual objects. This leaves open the possibility that virtual fictionalists like McDonnell and Wildman would acknowledge that virtual objects can casually impact conventional physical

objects. Insofar as this is the case and hologrammatically extended cognition *does not* require causal interactions between virtual objects (but only causal interactions between virtual objects and conventional physical objects), then virtual fictionalism might be consistent with hologrammatically extended cognition.

As should be evident, then, whether the HoloFoldit case represents an instance of hologrammatically extended cognition, and whether hologrammatically extended cognition is possible in the first place, hinges upon nuanced issues in the metaphysics of virtual reality (only a few of which have been mentioned here). Smart's paper beautifully lays the groundwork for further research into how emerging mixed reality applications facilitate novel forms of extended cognition.

Acknowledgements I am immensely grateful to the Notre Dame-IBM Technology Ethics Lab for a research fellowship that has granted me the time and funding necessary to complete this commentary article.

Author Contribution Not applicable.

Funding This work is supported by the Notre Dame-IBM Tech Ethics Lab as part of a postdoctoral fellowship.

Data Availability Not applicable.

Declarations

Ethics Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication The author hereby consents to the publication of this work.

Competing Interests The author declares no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Chalmers, D. J. (2017). *The virtual and the real*. *Disputatio*, 9(46), 309–352.
- Chalmers, D. J. (2022). *Reality+: Virtual Worlds and the Problems of Philosophy*. W. W. Norton.
- Clark, A., & Chalmers, D. J. (1998). The extended mind. *Analysis*, 58(1), 7–19.
- Logg, Anders et al. (2017). 'Solving Poisson's equation on the Microsoft HoloLens.' In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, ACM, New York, NY, USA, Article 87.
- McDonnell, N., & Wildman, N. (2019). Virtual reality: Digital or fictional? *Disputatio*, 11(55), 371–397.

- Palermos, S. O. (2014). Loops, constitution, and cognitive extension. *Cognitive Systems Research*, 27, 25–41.
- Rupert, R. D. (2004). Challenges to the hypothesis of extended cognition. *Journal of Philosophy*, 101(8), 389–428.
- Smart, P. (2018). Human-extended machine cognition. *Cognitive Systems Research*, 49, 9–23.
- Smart, P. (2022). Minds in the metaverse: Extended cognition meets mixed reality. *Philosophy and Technology*, 35(4), 1–29.
- Smart, P., & Clowes, R. (2021). Intellectual virtues and Internet-extended knowledge. *Social Epistemology Review and Reply Collective*, 10(1), 7–21.
- Smart, P. R. (2020). The Joi of holograms. In T. Shanahan & P. R. Smart (Eds.), *Blade runner 2049: A philosophical exploration* (pp. 127–148). Routledge.
- Smart, P. R. (2017). Extended cognition and the internet: A review of current issues and controversies. *Philosophy & Technology*, 30(3), 357–390.
- Walton, K. L. (1990). *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Harvard University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.