



Whose (Extended) Mind Is It, Anyway?

Keith Harris¹ 

Received: 6 October 2018 / Accepted: 16 September 2019 / Published online: 20 September 2019
© Springer Nature B.V. 2019

Abstract

Presentations of the extended mind thesis are often ambiguous between two versions of that thesis. According to the first, the extension of mind consists in the supervenience base of human individuals' mental states extending beyond the skull and into artifacts in the outside world. According to a second interpretation, human individuals sometimes participate in broader cognitive systems that are themselves the subjects of extended mental states. This ambiguity, I suggest, contributes to several of the most serious criticisms of the extended mind thesis, for these criticisms only apply to the first interpretation of the thesis. In what follows, I argue that several significant objections to the extended mind thesis fail to undermine the latter interpretation of that thesis. Having defended the second interpretation, I argue that the extension of mind does not involve the extension of self. Consequently, the subject of extended mental states is not the same individual whose causal coupling with external artifacts gives rise to extended mentality.

1 Introduction

Since it first came to prominence in philosophical thought, the extended mind thesis has animated sustained and vigorous debate. Yet while considerable attention, both critical and supportive, has been devoted to the thesis, certain fundamental issues are routinely overlooked. Presentations of the extended mind thesis are often ambiguous between two versions of that thesis. According to the first, the extension of mind consists in the supervenience base of human individuals' mental states extending beyond the skull and into artifacts in the outside world. According to a second interpretation, human individuals sometimes participate in broader cognitive systems that are themselves the subjects of extended mental states. While this ambiguity has not gone unnoticed (Adams and Aizawa 2008), its significance is often underappreciated.

✉ Keith Harris
keithraymondharris@gmail.com

¹ University of Iowa, Iowa City, IA 52245, USA

This ambiguity, I suggest, contributes to several of the most serious criticisms of the extended mind thesis, for these criticisms undermine only the first interpretation of the thesis. In what follows, I argue that several significant objections to the extended mind thesis fail to undermine the latter interpretation of that thesis. Having defended the second interpretation, I argue that the extension of mind does not involve the extension of self. Consequently, the subject of extended mental states is not the same individual whose causal coupling with external artifacts gives rise to extended mentality.

2 The Extended Mind

Andy Clark and David Chalmers (1998) are largely responsible for bringing the extended mind thesis (EMT) to philosophical prominence. Clark and Chalmers ask us to imagine Otto, whose Alzheimer's disease has deprived him of long-term memory capacities (1998, pp. 12–13). To compensate, Otto records new information in a notebook that he carries with him consistently. When Otto needs a piece of information, he recovers this information from his notebook. When, for instance, Otto desires to visit the Museum of Modern Art, he consults his notebook to recall its location and thereby comes to occurrently believe that the museum is located on 53rd Street—call this proposition *m*. Here it is useful to contrast Otto with Inga, who suffers no significant deficiencies in memory. We may suppose that Inga, like Otto, wishes to visit the Museum of Modern Art. Inga remembers the location of the museum, but this information is stored¹ in her long-term memory and is not occurrent to her. Recalling the location of the museum requires that Inga consult her long-term memory.

The simplest case for the extended mind thesis is built on the apparent equivalence between Otto and Inga. Inga, it seems, believes that *m* before sifting through her long term-memory and thereby making that belief occurrent. To deny this would do violence to our ordinary conceptions of the range of beliefs had by typical believers. Yet the role that Inga's long-term memory plays for her is strikingly like the role that Otto's notebook plays for him. The core difference between the two is that, while Inga stores information internally, Otto stores information in a physical artifact. Yet to insist, on this basis, that Inga and not Otto has a belief about the location of the museum, would betray a sort of biologically-centered prejudice (Clark 2008).

¹ One might reasonably object to my presentation of the Inga case that it reflects an objectionable commitment to a storehouse model of biological memory. This passive conception of memory has largely been abandoned in favor of an active or generative model of memory, and I write in terms of information storage here only for ease of exposition. Notably, once we acknowledge that biological memory is active, one might be concerned that we have lost the supposed parity between Otto and Inga, as Otto's notebook *is* a passive store of information. Clark has addressed this objection elsewhere (2010a, pp. 52–53), arguing that while there is a difference between Otto's memory system and Inga's this difference is not relevant. Clark's contention is that, because we would not hesitate to think of a passive internal system as part of a cognizer's cognitive apparatus, the active–passive distinction provides no basis for denying that Otto's notebook is a partial realizer of his mind.

The case-based argument for the extended mind thesis trades on a ‘common-sense functionalism’² about belief. Clark and Chalmers supplement the case-centered argument with an appeal to the following principle:

Parity Principle

If, as we confront some task, a part of the world functions as a process which, were it to go on in the head, we would have no hesitation in accepting as part of the cognitive process, then that part of the world is (for that time) part of the cognitive process. (Clark and Chalmers 1998, p. 8)

The parity principle explicitly denies that the boundaries of the skull mark a theoretically significant border between the cognitive and the non-cognitive. Although the principle explicitly concerns cognitive *processes*, similar considerations of parity militate against a conception of mental states that precludes their partial realization in material artifacts.

In addition to the direct intuitive appeal of the case and their appeal to the parity principle, Clark and Chalmers offer a series of considerations that make it plausible that the supervenience base of Otto’s belief that *m* includes his notebook (1998, p. 17):

Constancy

Otto carries the notebook with him consistently, and rarely acts without consulting it.

Availability

Otto can access the information in his notebook without difficulty.

Automatic Endorsement

When Otto retrieves information from his notebook, he accepts it without scrutiny.

Prior Endorsement

Otto has consciously endorsed the information contained in the notebook at some point in the past, and the information is recorded there *because* he has previously endorsed it.

Clark and Chalmers do not intend constancy, availability, automatic endorsement, and prior endorsement as models of necessary and sufficient conditions on the integration of artifacts in extended cognitive systems. Indeed, the authors express

² Clark (2008, p. 88) adopts this terminology from David Braddon-Mitchell and Frank Jackson (2007, Ch. 3). Common-sense, or analytic, functionalism stands in contrast to ‘empirical’ or ‘psychofunctionalist’ variants of functionalism that characterize mental states narrowly in the language of neurophysiology. One cost of this latter branch of functionalism is the inescapable denial that beings with physiologies substantially different to ours can have mental states like ours (Braddon-Mitchell and Jackson 2007, p. 88). Consequently, this latter branch of functionalism precludes extended mental states from the outset.

immediate reservations about the relevance of prior endorsement to extended belief (1998, p. 17) and Clark has since reiterated concerns about its relevance (2008, p. 80). Still, the four considerations are suggestive of the circumstances under which an artifact is a candidate for inclusion in an extended mental system. As we will see in what follows, critics of the extended mind thesis have often drawn attention to difficulties that arise concerning these four considerations.

3 Objections to the Extended Mind Thesis

The extended mind thesis has provoked sustained criticism. Here my focus is on three prominent lines of objection against the thesis. The objections to the extended mind thesis considered here by no means exhaust the objections previously levelled against that thesis. In addition to the objections mentioned below, sustained criticisms of the thesis are offered by, for instance, Frederick Adams and Kenneth Aizawa (2001, 2008) and Robert Rupert (2009). Clark (2008, 2010a, b) has previously offered convincing responses to several of these lines of objection, and so I focus here on a subset of objections that are either persistent or recent. In addressing these objections, I will not attempt to show that the extended mind thesis is beyond criticism. Rather, my aim is to show that, understood correctly, the extended mind thesis is at least more resilient to objection than is often supposed.

3.1 The Otto Two-Step

The extended mind thesis is a radical departure from a certain commonsense—or, at least, commonly-held—conception of the mind and its relation to the world. It is thus unsurprising that opponents of the extended mind thesis have proposed alternative, relatively conservative, accounts of what occurs in supposed cases of extended mentality. Such accounts retain commitment to the orthodox view that individuals' mental states are organismically-bounded. Given the relative simplicity of orthodox views (Rupert 2009, p. 46, 50), the availability of such accounts significantly undercuts the motivation for positing extended mental states.

Clark and Chalmers anticipate one such response, which, applied to the Otto case, Clark (2010a) has dubbed the 'Otto two-step.' In that case, the concern is that Otto's behavior can be fully accounted for without allowing that Otto believes that *m* prior to consulting his notebook. One may instead suppose that Otto believes that the location of the museum is recorded in his notebook and, in light of this belief, consults his notebook to learn the location of the museum. Only following this process does Otto himself believe that *m* (Clark 2010a, p. 46). Or so the objection goes.

The Otto two-step is bolstered by some features of the Otto case that are puzzling on the supposition that Otto has a belief about the location of the museum prior to consulting his notebook. First, it is seemingly not until Otto perceives the location of the museum, as recorded in his notebook, that he comes to have an occurrent belief as to its location. In this respect, Otto is significantly different to Inga, whose recollection of the location of the museum is clearly not mediated by perception (Butler

1998, p. 211). A second and related line of support for the Otto two-step centers on the fact that, to recall the location of the museum, Otto first checks the notebook, that is, Otto *does something* as a step toward determining the location of the museum (Wikforss 2014, p. 470). Yet, on the supposition that Otto already believes that *m*, consulting the notebook seems to be, from Otto's perspective, a wasted effort.

Clark (2010a) has a ready response to the Otto two-step and related concerns. If Otto's relation to the notebook was more like a cognitively normal individual's relation to a phonebook, say, then Otto's consultation of the notebook would be a proper action and his belief would be formed by the process, rather than merely recalled. Yet Otto's relation to the notebook is not like this. The supervenience base of certain of Otto's mental states, including his belief that *m*, includes the notebook. Because Otto is related to the notebook in this way, his consultation does not involve perception, but something more akin to introspection (Clark 2010a, p. 56). Similarly, given Otto's connection to the notebook, his consultation of the notebook is not a proper action, but is instead an instance of thinking (Clark 2010a, p. 47). Thus, Clark (2010a) maintains that, properly understood, Otto's consultation of the notebook does not objectionably involve perception or action any more so than Inga's process of recollection (2010a, p. 46).

The success of Clark's response to the Otto two-step is, I suggest, unclear. Clark offers one response to the objection that the proponent of the extended mind thesis can comfortably accept, yet this response is unlikely to convince those not already convinced by the thesis. For one thing, the thought that Otto is *doing something* and not merely thinking when he consults the notebook is difficult to shake. I suggest in Sect. 4 that one interpretation of the extended mind thesis can account for such reservations while also accounting for what Clark's response to the Otto two-step gets right.

3.2 Cognitive Bloat

Among the features of the Otto case that render it plausible to ascribe to Otto the belief that *m* is, according to Clark and Chalmers, the fact that Otto has previously endorsed this proposition. Indeed, Otto's notebook contains this information *because* Otto has previously endorsed it. Clark and Chalmers (1998) are quite explicit about their uncertainty as to the significance of prior endorsement. As Clark makes clear, consideration of prior endorsement raises the following dilemma for proponents of the extended mind thesis. On the one hand, prior endorsement seems an implausibly strict restriction on extended belief insofar as ordinary beliefs are regularly formed through unconscious processes. However, in the absence of some such condition on extended belief, coupling with certain varieties of external artifacts will generate an implausibly vast stock of extended beliefs (Clark 2008, p. 80). It is worth expanding on this concern below so as to make the dilemma faced by the proponents of the extended mind thesis clear.

As I noted above, the basic case for the extended mind thesis relies on a common-sense functionalism about mental states. On this view, what makes something a belief is its functional role in its possessor's mental operations. For this reason, any

difference in functional role between ordinary beliefs and extended beliefs undercuts the motivation for understanding the latter as proper beliefs. But many of our ordinary beliefs have never been consciously endorsed. Thus, the inclusion of prior endorsement as a condition on extended belief threatens to sever functional parity between ordinary and extended beliefs, thereby undercutting the basis for countenancing the latter.

The proponent of extended belief cannot avoid the present worry by simply abandoning prior endorsement as a condition on extended belief. The removal of this condition effectively opens the floodgates to the number of extended beliefs to which proponents of extended belief are committed. The worry here is about ‘cognitive bloat’³ (Clark 2008; Wikforss 2014). Suppose, for instance, that an individual consistently carries a local phonebook with him, never making a local call without looking up the number in the phonebook. Whereas an extended mind thesis that includes prior endorsement as a condition on extended belief could deny that this individual has beliefs about all listed local numbers, the abandonment of this condition seems to saddle the proponent of extended beliefs with the undesirable commitment to the imagined individual having an implausibly vast stock of beliefs concerning local numbers.

EMT thus produces a dilemma. That thesis either suggests the existence of an implausibly vast number of extended beliefs or places a condition on extended belief that reduces its functional parity with ordinary belief, thereby undercutting the functionalist grounds for accepting the thesis. One possibility is that some alternative condition might serve, like prior endorsement, to significantly cut down on the number of extended beliefs to which proponents of the extended mind thesis must commit. I set this possibility aside here. Instead, I show that there is a construal of the extended mind thesis on which cognitive bloat is no threat to the thesis.

3.3 Extended Knowledge

Aizawa (2018) suggests that, not only do proponents of the extended mind run the risk of attributing implausibly many beliefs to certain individuals, they are likewise committed to certain individuals possessing an implausibly vast stock of knowledge. Aizawa’s concern is a response to a recent trend among certain epistemologists and philosophers of mind to countenance extended knowledge, in addition to extended belief.⁴ Aizawa presents the case of Otis, an organic chemistry student who, in lieu

³ While Clark and Åsa Wikforss seem to understand ‘cognitive bloat’ as the concern that endorsement of EMT commits one to the attributing of an implausibly vast range of beliefs to certain individuals, Sean Allen-Hermanson (2013) uses the term to designate the concern that EMT entails an implausible expansion of the mind such that its supervenience base includes physical objects that do not plausibly bear that relation to the mind. I focus here on the former version of the concern, but it should be clear that the revised interpretation of EMT I advocate is more resilient to both versions of the concern than the orthodox interpretation.

⁴ Early contributors to the literature on extended knowledge include Pritchard (2010), Kelp (2013), Berit Brogaard (2014), S.O. Palermos (2014), and Clark (2015). For a recent collection of relevant work in this area, see J.A. Carter et al. (2018).

of studying for his exam, simply copies information from his textbook onto notecards, which he consults during the exam (2018, p. 74). Arguably, the role the notecards play for Otis is functionally indistinguishable from the role that a more dedicated student's brainbound memories play for her. Nonetheless, Aizawa suggests plausibly⁵ that Otis does not have a respectable level of knowledge about organic chemistry. It bears noting that the present objection is not merely a rehashing of the dilemma described in the preceding subsection, as the intuition that Otis fails to know what is written on his notecards persists even if we allow that Otis has previously consciously endorsed the content of those cards.

A proponent of the extended mind thesis might respond that Otis's relationship to the notecards is not sufficiently like Otto's relationship to the notebook for the information written on the notecards to contribute to his stock of beliefs or his stock of knowledge. It is difficult to imagine, for instance, that the notecards are a constant in Otis's life. Such a response would miss the thrust of the objection, however. We can easily imagine an inveterate cheater with a long history of using the same notebook to cheat on his exams. We may even suppose that this individual recognizes the relevance of the information contained in his notebook outside of exam scenarios and uses this information where appropriate. Still, it remains implausible that such an individual knows what is written in his notebook. I now turn to a disambiguation of the extended mind thesis that yields two distinct versions of that thesis, one of which avoids attributing the contested knowledge to Otis and other apparent implausible implications of the view.

4 Two Versions of the Extended Mind Thesis

Discussion of EMT is often clouded by a central and significant ambiguity. To introduce this ambiguity, it is useful to compare it to a related ambiguity concerning extended cognition. The extended cognition thesis, according to Adams and Aizawa, is ambiguous between positing *extended cognitive systems* and *extended cognitive processes*. To posit extended cognitive systems is to suggest that individuals' brains sometimes function as parts of broader cognitive systems that implement cognitive processes not attributable to the individuals alone. In contrast, extended cognitive processes are cognitive processes carried out by brain, body, and environment that are nonetheless attributable to the individual whose brain participates in these processes.

The related ambiguity we are interested in here concerns the subject of extended mental states. Consider two interpretations of the Otto case. According to the first,

⁵ The attribution of a significant range of knowledge of organic chemistry to Otto is a bullet that some proponents of EMT are likely willing to bite. While I will not defend Aizawa's counterexample at length here, it is worth pointing out the broader point of which that counterexample is illustrative. As Aizawa illustrates, acceptance of EMT and the possibility of extended knowledge has severely revisionist implications for the accuracy of our knowledge-attributing practices. Other things being equal, these implications put at least some pressure on EMT, as it is typically understood. Thanks to an anonymous reviewer for pressing me on this point.

Otto's relationship to the notebook is such that the supervenience base for his belief states includes the contents of the notebook. According to a second interpretation, Otto's relationship to the notebook is such that the notebook makes up part of the supervenience base for the mental states of a distinct agent of which both Otto and the notebook are parts—call this Otto+.

The former interpretation of the extended mind thesis is often explicitly defended by or, more often, suggested by proponents of that thesis. Thus, Clark and Chalmers (1998) and, in later individual work Clark himself (2008, 2010a), consistently write that the supervenience base of *Otto's* beliefs extends beyond his skull. Parallels draw between Otto and Inga suggest the same interpretation, as there is little temptation to suppose that the subject of the belief that *m*, before Inga consults her long-term memory, is a macro-agent composed of Inga, together with her long-term memory. More generally, Clark writes that proponents of the extended mind thesis:

[P]aint the mind itself (or better, the physical machinery that realizes some of *our cognitive processes and mental states*) as, under humanly attainable conditions, extending beyond the bounds of skin and skull. (2008, p. 76, *italics added*)

Robert Wilson suggests this former interpretation more-or-less explicitly, writing that:

[I]n both species of externalism that I have discussed, the individual remains the subject or bearer of psychological states, even if she no longer serves as a boundary demarcating the entities of a respectable psychological science. (2004, pp. 212–213)

More generally, Wilson maintains that, although individuals' mental states are often partially realized in features of the individuals' social and physical environments, those mental states are nonetheless properties of the individuals in question (2004, p. 141). For the sake of brevity, let us call this view EM-1.⁶

The latter interpretation of the extended mind thesis, according to which the subjects of extended mental states are systems consisting of individuals in conjunction with material artifacts, has been developed in some detail by Yasuo Nakayama (2013). Call this view EM-2. It has elsewhere been noted that EM-2 is closely analogous to the so-called 'systems reply' to John Searle's (1980) Chinese Room thought experiment (Miyazono 2017). According to the systems reply to Searle's thought experiment, Searle is correct to think that he—equipped with a stock of Chinese symbols, rules in English for correlating these, and a set of instructions in English for doing so—understands Chinese, but is wrong to conclude that no understanding of Chinese is present in the scenario. While Searle himself fails to understand Chinese, the broader system of which he is a part understands Chinese. Similarly, on

⁶ The view espoused in Wilson (2004) has much in common with the 'wide computationalist' view Wilson explicates in his (1994) paper of the same name. In this earlier work, Wilson does not explicitly commit to the attribution mental states with a wide realization base to biological individuals or the systems in which they are embedded.

EM-2, Otto fails to possess certain mental states that are possessed by the broader system of which he is a part. Thus, insofar as functionalists have reason to adopt the systems reply in response to what would otherwise be a troubling objection, functionalists have some motivation to accept EM-2, instead of EM-1.

Although the primary proponents of the extended mind thesis often suggest acceptance of EM-1, they sometimes suggest sympathy with EM-2. Clark (2010b, pp. 96–97), for instance, writes that, when searching for the locus of cognition, “we must attend...to the larger organization of which biological Otto is a part.” The apparently explicit endorsement of EM-1 by the primary proponents of the extended mind thesis can be reconciled with their suggestions of the latter interpretation by proper attention to boundaries of the self. Clark and Chalmers write that “Otto *himself* is best regarded as an extended system, a coupling of biological organism and external resources” (1998, p. 18). The view espoused here seems to be that Otto is not himself a biological organism, but is instead a hybrid composed of biological Otto and his notebook. Thus, where Clark and Chalmers use ‘Otto,’ the intended referent is apparently the agent I have called ‘Otto+.’

Clark and Chalmers’s apparent support for EM-1 is therefore misleading. Although the authors attribute extended mental states to Otto, they regard Otto himself as consisting of a biological individual in conjunction with a material artifact. Thus, Clark and Chalmers appear to endorse EM-2. I suspect that the ambiguity of the extended mind thesis is overlooked by many critics of that thesis for, as we will see, some of the more serious objections to that thesis have no traction against EM-2.

5 Dissolving Objections to the Extended Mind Thesis

In Sect. 2, I considered three objections to the extended mind thesis. I now argue that, although these objections make trouble for EM-1, they pose no challenge to EM-2.

5.1 The Otto Two-Step Revisited

As we have seen, one tempting response to the Otto case is to suppose that, prior to consulting his notebook, Otto believes only that the location of the Museum of Modern Art is recorded in his notebook. According to this response, Otto learns that *m* by consulting his notebook, rather than merely making his belief that *m* occur-rent. Proponents of the extended mind thesis respond that, given his connection with the notebook, Otto’s consultation involves introspection, rather than action (Clark 2010a, p. 57). As we have seen, the case is such that proponents and opponents of the extended mind thesis alike run the risk of begging the question against their counterparts.

Yet the apparent plausibility of each position is, I suggest, due to an ambiguity as to the referent of ‘Otto.’ As we have seen, some proponents of the extended mind thesis take ‘Otto’ to refer to a hybrid entity consisting of the biological Otto,

together with his notebook. On this picture, the Otto two-step is no more plausible than a parallel interpretation of the Inga case.

By contrast, detractors of the extended mind thesis seem to take 'Otto' to refer strictly to a biological organism, one coupled with, but not partly constituted by, a material artifact. Consequently, these opponents of the thesis think Otto's retrieval of information concerning the location of the museum involves acting on an object external to himself, a process involving perception. On this view, the two-step interpretation of the Otto case seems highly plausible.

The suggestion sketched here is that the Otto two-step rests on a misunderstanding. Properly conceived, the retrieval of the information that *m* is carried out by an individual, Otto+, partially constituted by a material object that itself forms part of the supervenience base for that individual's beliefs. Thus, the process involves something akin to introspection, not perception, and resembles thinking rather than acting.

5.2 Cognitive Bloat Revisited

Recall that the proponent of extended belief apparently faces a dilemma concerning prior endorsement as a condition on extended belief. Inclusion of such a condition threatens to undermine functional parity between extended and ordinary belief, while exclusion of this condition opens the door to cognitive bloat. I now argue that this dilemma can be dissolved by denial of EM-1 in favor of EM-2.

As we have seen, the inclusion of prior endorsement as a condition on extended belief reduces the functional parity that obtains between ordinary belief and extended belief, thereby undercutting the basis for postulating the latter. Fortunately, the proponent of extended belief may deny prior endorsement as a condition on extended belief without being saddled with the unwelcome implication of cognitive bloat. Here it is vital to understand why cognitive bloat is typically taken to be problematic. What is clear, I take it, is that one cannot vastly increase one's stock of beliefs simply by carrying around, and consistently relying upon, a phonebook, smartphone, or some such artifact. However, it is perfectly consistent with the accuracy of this intuition that, in carrying around such an artifact, one participates in a broader cognitive system whose beliefs supervene, in part, on the material artifact. Just as it is widely acknowledged that a group may know vastly more than its individual members (Bird 2010, 2014; Lackey 2014), it is likewise plausible that individuals may participate in extended cognitive systems that believe more than themselves.

5.3 Extended Knowledge Revisited

Related to the worry about cognitive bloat is a concern that the extended mind thesis, in conjunction with plausible assumptions concerning the generation of knowledge, will improperly attribute knowledge to certain individuals. This worry fails to undermine the extended mind thesis for much the same reason as the preceding worry about cognitive bloat. To see this, let us return to the case of Otis. As we saw,

it is implausible to attribute to Otis knowledge of all the information written on his notecards. However, it does not follow that it is implausible to allow that there is an individual distinct from Otis but of which Otis is a part—call this ‘Otis+’—that possesses the knowledge in question. We can respect our intuition that Otis fails to know much about organic chemistry, and for this reason does not deserve to do well on his exam, by denying the identity of Otis and Otis+. On EM-2, we may allow that the latter possesses beliefs and knowledge that the former does not.

6 Extended Minds and Extended Selves

If we allow that the subjects of extended mental states are hybrid individuals composed of biological individuals together with physical artifacts, questions remain concerning the identity of the hybrid individuals in question. As we have seen, Clark and Chalmers suggest that, when the mind extends, so too does the self. For instance, it is not simply the supervenience base of Otto’s mental states that includes material artifacts, but also the supervenience base of Otto himself. I now argue that this view is mistaken. The subject of extended mental states is not Otto, but the distinct agent Otto+. More generally, individual selves are not extended in cases of extended mentality.⁷ Rather, cases of extended mentality involve the generation of new agents, psychologically continuous with, but distinct from, the human individuals that partly compose them.

I wish to begin by addressing one objection to a proposal that is similar, in relevant ways, to the one made here. Kengo Miyazono (2017) argues that proponents of the extended mind thesis fail to justify the claim that extended mental states are states of human individuals. In response, Miljana Milojevic (2018) argues that, according to a plausible criterion of personal identity, human individuals are identical to the extended individuals they give rise to by coupling with external artifacts. Milojevic appeals to the following criterion of identity, suggested by Derek Parfit (1984, p. 207):

The Psychological Criterion

(1) There is *psychological continuity* if and only if there are overlapping chains of strong connectedness. X today is one and the same person as Y at some past time if and only if (2) X is psychologically continuous with Y, (3) this continuity has the right kind of cause, and (4) it has not taken a ‘branching’ form. (5) Personal identity over time just consists in the holding of facts like (2) and (4).

⁷ It bears noting that there exists a small body of literature on the prospects for extended selfhood in light of the extended mind thesis. These discussions draw on contentious claims regarding the significance of consciousness (Buller 2013), memory (Heersmink 2017), durability (Baker 2009), and organismic boundaries (Olsen 2011) for selfhood. Here I adopt a different tactic, arguing that the view that the self extends has implausible consequences in the Otis case, conflicts with a close theoretical relative of the extended cognition thesis, and mischaracterizes the relationship between Otto’s condition and his notebook. This strategy has the advantage of requiring no contentious general assumptions about the boundaries of the self.

Milojevic's case for the extension of self turns on the fact, which I do not intend to deny, that extended agents are psychologically continuous with purely human agents. Even granting this continuity and granting the adequacy of the psychological criterion, however, we lack grounds for identifying extended agents with the purely human agents whose coupling with external artifacts gave rise to the extended agents. On the view I am developing here, and that Miyazano's suggests, there are two individuals psychologically continuous with biological Otto who, at time, t_1 , becomes coupled to the notebook in the way Clark and Chalmers describe. At subsequent time, t_2 , two distinct agents, biological Otto and Otto+, are psychologically continuous with biological Otto at t_1 . Although different to the 'branching' cases originally sketched by Parfit, cases of extended mind involve, on the present account, a branching of psychological continuity between a purely biological and an extended agent. Hence appeal to the psychological criterion is, in virtue of (4), not enough to secure the identity of Otto+ and Otto prior to t_1 .

Moreover, we have already seen one reason to prefer the interpretation of EM-2 according to which the extension of mind does not involve the extension of self. In the Otis case described above, EM-2 avoids the implausible attribution of knowledge to Otis, but only if we deny that Otis's mind extends to include, as part of its supervenience base, his notecards. If it is *Otis's* mind that extends, then Otis is indeed the subject of the knowledge in question. But this is precisely the implication that the move from EM-1 to EM-2 is meant to avoid.

A second reason to deny that the extension of mind involves the extension of self comes from consideration of distributed cognition. Like the extended cognition thesis, the distributed cognition thesis has it that cognitive processes are not confined to biological individuals. However, whereas the extended cognition thesis has it that cognition takes place between single individuals and material artifacts, the distributed cognition thesis has it that cognition may be distributed among several individuals and material artifacts. A classic and illustrative study in distributed cognition is Edwin Hutchins's (1995) work on cognition among the navigation team of a ship Hutchins gives the pseudonym 'U.S.S. Palau'. The crucial point for our purposes is that Hutchins, as well as later commentators, attributes distributed cognitive processes and mental states to the entire navigation team, rather than individuals within that team. In such cases, individuals, along with material artifacts, form part of the supervenience base for cognitive processes and mental states of the team itself. Yet, in such cases, there is no temptation to suppose that individuals themselves extend to encompass the team. While individual crewmembers of the U.S.S. Palau participate in a complex system of causal interactions with other crewmembers and with material artifacts aboard the ship, it would be implausible to identify any individual with this system. Instead, individuals maintain their identities while participating in broader cognitive systems. The distributed cognition thesis is not typically taken to be incompatible with the extended cognition thesis. Indeed, we might understand extended cognition as a limiting case of distributed cognition, where just one individual participates in a distributed cognitive system. Thus, if we do not allow that individual selves are extended in the case of distributed cognition, it seems there is no reason to allow that individual selves are extended in the case of extended cognition.

Finally, while the extended mind thesis highlights the functional parity between certain material artifacts and more traditional objects of cognitive science, this parity ought not be exaggerated. Recall that Otto records information in his notebook because his Alzheimer's disease has significantly reduced his ability to rely on his biological memory. If the extended mind is attended by the extension of self, Otto's notebook then serves as a physical substratum of *his* memory, in much the same way that part of Inga's brain does for her. On this view, once Otto begins using the notebook, he faces no significant deficiency in memory. But this is implausible. The use of a notebook does not cure Otto's Alzheimer's disease, the notebook merely helps to compensate for it (Preston 2010, p. 365). That Otto's deficiency in memory persists despite his use of the notebook is consistent with there being a second agent—Otto+—that suffers no deficiency in memory.

7 Concluding Remarks

To conclude, I wish to address two further objections to the proposal advanced here. First, one might be concerned that, while EM-2 avoids problems faced by EM-1, this revised understanding of EMT, together with the claim that extended agents are not numerically identical to the biological agents whose coupling with external artifacts brings them about, entails an overabundance of hybrid agents.⁸ Many of us, for instance, depend heavily on the use of one or more smartphones, a work computer, a home computer, and a broad range of other instruments. Given these various relationships, one might be concerned that commitment to EM-2 demands commitment to our partial constitution of various hybrid entities distinct from ourselves.

The proponent of EM-2 has several strategies available for responding to the present line of objection. First, although we rely on a wide range of objects external to ourselves, we are connected to few of these objects as intimately as Otto is connected to his notebook. Even for those who rely heavily on their home and work computers, for instance, it is unclear that this dependency is sufficient to secure the *availability* and *constancy* conditions described in Sect. 1. Thus, the proponent of EM-2 can to some degree resist the overgeneration of hybrid agents by invoking these conditions. Still, a defense of EM-2 cannot rest on this straightforward reply because (1) as I noted above, Clark and Chalmers (1998) do not construe *availability* and *constancy* as necessary conditions on some object partially constituting the supervenience base of one's mental states and (2) even interpreting *availability* and *constancy* as necessary conditions might fail to rule out overgeneration of hybrid agents. Supposing my relationships to my work smartphone and personal smartphone satisfy *availability* and *constancy*, do I now partially constitute at least two distinct extended agents? The proponent of EM-2 need not answer affirmatively. Nothing in EM-2 commits one to the existence of a distinct hybrid agent for every relationship that satisfies *availability*, *constancy*, and, *automatic endorsement*. One might, consistently with EM-2, maintain that in cases where multiple such relationships obtain there is a

⁸ Thanks to an anonymous reviewer for raising this important objection.

single hybrid agent whose mental states supervene on states of the biological individual and all the artifacts to which that individual is coupled.

Second, one might worry that interpreting the extended mind thesis as EM-2, while denying the extension of the self, salvages the plausibility, but not the novelty, of the extended mind thesis.⁹ Indeed, on the proposal advanced here, ‘extended mind’ is a misnomer, as no individual’s mind extends into the environment. In reply, we should notice, first, that whether a theory is interesting is no marker of truth. Second, even if some of the revisionary implications of the extended mind thesis are nullified by the present account, not all are. If EM-2 is true and the self does not extend, one cannot expect to extend one’s mind by coupling with material artifacts. One can, however, expect to encounter agents that are composed of biological individuals in conjunction with material artifacts. Thus, the present account arguably loses some of the novelty of the extended mind thesis as conceived from the first-person perspective,¹⁰ while retaining its novelty from the third-person perspective. These revisionist implications are enough to open new lines of inquiry in epistemology, ethics, and other domains.

References

- Adams, F., & Aizawa, K. (2001). The bounds of cognition. *Philosophical Psychology*, *14*(1), 43–64.
- Adams, F., & Aizawa, K. (2008). *The bounds of cognition*. Malden, MA: Wiley-Blackwell.
- Aizawa, K. (2018). Extended cognition, trust and glue, and knowledge. In J. A. Carter, A. Clark, J. Kallestrup, S. O. Palermos, & D. Pritchard (Eds.), *Extended epistemology* (pp. 64–78). Oxford: Oxford University Press.
- Allen-Hermanson, S. (2013). Superdupersizing the mind: Extended cognition and the persistence of cognitive bloat. *Philosophical Studies*, *164*, 791–806. <https://doi.org/10.1007/s11098-012-9914-7>.
- Baker, L. R. (2009). Persons and the extended mind thesis. *Zygon*, *44*(3), 642–658. <https://doi.org/10.1111/j.1467-9744.2009.01022.x>.
- Bird, A. (2010). Social knowing: The social sense of ‘scientific knowledge’. *Philosophical Perspectives*, *24*(1), 23–56. <https://doi.org/10.1111/j.1520-8583.2010.00184.x>.
- Bird, A. (2014). When is there a group that knows? Distributed cognition, scientific knowledge, and the social epistemic subject. In J. Lackey (Ed.), *Essays in collective epistemology* (pp. 42–63). Oxford: Oxford University Press.
- Braddon-Mitchell, D., & Jackson, F. (2007). *Philosophy of mind and cognition: An introduction* (2nd ed.). Oxford: Blackwell Publishing.
- Brogaard, B. (2014). A partial defense of extended knowledge. *Philosophical Issues*, *24*(1), 39–62.
- Buller, T. (2013). Neurotechnology, invasiveness and the extended mind. *Neuroethics*, *6*(3), 593–605. <https://doi.org/10.1007/s12152-011-9133-5>.
- Butler, K. (1998). *Internal affairs: A critique of externalism in the philosophy of mind*. Dordrecht: Kluwer.
- Carter, J. A., Clark, A., Kallestrup, J., Palermos, S. O., & Pritchard, D. (Eds.). (2018). *Extended Epistemology*. Oxford: Oxford University Press.

⁹ Milojevic (2018), for instance, suggests that Miyazono’s (2017) application of the systems reply to the extended mind thesis, if successful, renders the thesis uninteresting.

¹⁰ The magnitude of this loss is unclear, however, as proponents of the extended mind thesis already acknowledge that, while some mental states are extended, it is unlikely that conscious ones are (Clark and Chalmers 1998; Clark 2008, p. 79).

- Clark, A. (2008). *Supersizing the Mind: Embodiment, action, and cognitive extension*. Oxford: Oxford University Press.
- Clark, A. (2010a). Memento's revenge: The extended mind extended. In R. Menary (Ed.), *The extended mind* (pp. 43–66). Cambridge: MIT Press.
- Clark, A. (2010b). Coupling, constitution, and the cognitive kind: A reply to Adams and Aizawa. In R. Menary (Ed.), *The extended mind* (pp. 43–66). Cambridge: MIT Press.
- Clark, A. (2015). What extended me knows. *Synthese*, 192, 3757–3775. <https://doi.org/10.1007/s11229-015-0719-z>.
- Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7–19. <https://doi.org/10.1093/analysis/58.1.7>.
- Heersmink, R. (2017). Distributed selves: Personal identity and extended memory systems. *Synthese*, 194, 3135–3151. <https://doi.org/10.1007/s11229-016-1102-4>.
- Hutchins, E. (1995). *Cognition in the wild*. Cambridge: MIT Press.
- Kelp, C. (2013). Extended cognition and robust virtue epistemology. *Erkenntnis*, 78, 245–252. <https://doi.org/10.1007/s10670-011-9301-3>.
- Lackey, J. (2014). Socially extended knowledge. *Philosophical Issues*, 24(1), 282–298. <https://doi.org/10.1111/phis.12034>.
- Milojevic, M. (2018). Extended mind, functionalism, and personal identity. *Synthese*. <https://doi.org/10.1007/s11229-018-1797-5>.
- Miyazono, K. (2017). Does functionalism entail extended mind? *Synthese*, 194, 3523–3541. <https://doi.org/10.1007/s11229-015-0971-2>.
- Nakayama, Y. (2013). The extended mind and the extended agent. *Procedia-Social and Behavioral Sciences*, 97, 503–510. <https://doi.org/10.1016/j.sbspro.2013.10.265>.
- Olsen, E. T. (2011). The extended self. *Minds and Machines*, 21(4), 481–495. <https://doi.org/10.1007/s11023-011-9258-7>.
- Palermos, S. O. (2014). Knowledge and cognitive integration. *Synthese*, 191(8), 1931–1951. <https://doi.org/10.1007/s11229-013-0383-0>.
- Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press.
- Preston, J. (2010). The extended mind, the concept of belief, and epistemic credit. In R. Menary (Ed.), *The extended mind* (pp. 355–369). Cambridge: MIT Press.
- Pritchard, D. (2010). Cognitive ability and the extended cognition thesis. *Synthese*, 175(1), 133–151. <https://doi.org/10.1007/s11229-010-9738-y>.
- Rupert, R. D. (2009). *Cognitive systems and the extended mind*. Oxford: Oxford University Press.
- Searle, J. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–457.
- Wikforss, Å. (2014). Extended belief and extended knowledge. *Philosophical Issues*, 24(1), 460–481. <https://doi.org/10.1111/phis.12043>.
- Wilson, R. A. (1994). Wide computationalism. *Mind*, 103(411), 351–372.
- Wilson, R. A. (2004). *Boundaries of the Mind: The individual in the fragile sciences*. Cambridge: Cambridge University Press.

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