



Phenomenal transparency and the extended mind

Paul R. Smart¹ · Gloria Andrada² · Robert W. Clowes²

Received: 24 March 2022 / Accepted: 21 July 2022

© The Author(s) 2022

Abstract

Proponents of the extended mind have suggested that phenomenal transparency may be important to the way we evaluate putative cases of cognitive extension. In particular, it has been suggested that in order for a bio-external resource to count as part of the machinery of the mind, it must qualify as a form of transparent equipment or transparent technology. The present paper challenges this claim. It also challenges the idea that phenomenological properties can be used to settle disputes regarding the constitutional (versus merely causal) status of bio-external resources in episodes of extended cognizing. Rather than regard phenomenal transparency as a criterion for cognitive extension, we suggest that transparency is a feature of situations that support the ascription of certain cognitive/mental dispositional properties to both ourselves and others. By directing attention to the forces and factors that motivate disposition ascriptions, we arrive at a clearer picture of the role of transparency in arguments for extended cognition and the extended mind. As it turns out, transparency is neither necessary nor sufficient for cognitive extension, but this does not mean that it is entirely irrelevant to our understanding of the circumstances in which episodes of extended cognizing are apt to arise.

Keywords Extended cognition · Extended mind · Phenomenal transparency · Mechanism · Disposition

✉ Paul R. Smart
ps02v@ecs.soton.ac.uk

Gloria Andrada
gandrada@fsh.unl.pt

Robert W. Clowes
robert.clowes@gmail.com

¹ Electronics & Computer Science, University of Southampton, Southampton SO17 1BJ, UK

² Instituto de Filosofia da Nova, Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa, Campus de Campolide - Colégio Almada Negreiros, 1099-032 Lisbon, Portugal

1 Introduction

According to proponents of the extended mind, “the physical machinery that realizes some of an individual agent’s cognitive processes and mental states can, under humanly attainable conditions, include elements and devices located beyond the bounds of skin and skull” (Clark, 2015, p. 3757).¹ This claim is sometimes made about explanatory kinds that are relevant to folk psychology (e.g., extended states of dispositional belief), while in other cases it is directed to explanatory kinds that are relevant to cognitive science (e.g., extended memory, extended problem-solving, etc.).² In both cases, however, there is a recognition that bio-external resources (e.g., a smartphone, a notebook, a mixed reality device, and so on) can become incorporated into an individual’s cognitive system, such that the resource in question is just as much a part of the machinery of the mind as is a biological brain region, such as the hippocampus or the visual cortex.

The question, of course, is what motivates this shift in the status of a bio-external resource? What is it that determines when some bio-external resource, such as a smartphone, makes the transition from a mere tool (something we use to scaffold our thinking) to a genuine mind part (something that forms the material fabric of our thinking)? In response to this question, it has been suggested that there is a certain phenomenology associated with cognitive extension. In particular, it has been suggested that in order for a resource to be a proper constituent of an individual’s cognitive system it should disappear from the conscious apprehension of the user, such that the user is no longer aware of the resource as an independent object (see Wheeler, 2019). The phenomenological property that is in play here is what we will call *phenomenal transparency*. It is the form of transparency³ that is commonly associated with philosophical discussions of tool use, particularly those arising from the work of phenomenological philosophers such as Heidegger (1927) and Merleau-Ponty (1945). Consider, for example, the case of the skilled carpenter who, while using a hammer, has no conscious apprehension of the hammer as an independent object. Rather than being aware of the hammer, the focus of attention for the carpenter is the task-at-hand—the hammer is no more the object of awareness for the carpenter than is the hand that wields the hammer. In such cases, the hammer is said to be transparent-in-use. It is ‘transparent’ in the sense that the carpenter ‘sees through’ the hammer to whatever carpentry-related task is being performed; they do not perceive the hammer as an independent object. The carpenter is, of course, performing a physical task, as opposed to a cognitive task; nevertheless, the same sort of idea underlies the appeal to transparency in an active externalist context. Once incorporated into a cognitive routine, bio-external resources are sometimes said to be a form of “transparent equipment” or “transparent technology” (Clark, 2003). They are transparent in the same

¹ See Clark and Chalmers (1998) and Clark (2008a) for canonical treatments of this idea.

² As a terminological note, we use the term “extended mind” to refer to situations involving folk psychological kinds (e.g., states of dispositional belief) and the term “extended cognition” to refer to situations involving cognitive scientific kinds (e.g., extended problem-solving). The term “cognitive extension” is used to refer to both extended cognition and the extended mind.

³ Phenomenal transparency is to be distinguished from other forms of transparency that have been discussed in the philosophical literature. For more on this, see Andrada et al. (2022).

way that the carpenter's hammer is transparent: In both cases, the user 'sees through' the bio-external resource to the (physical/cognitive) task at hand; they do not attend to the resource as an independent object, nor do they attend to the details of their interaction with the resource.

While phenomenal transparency is a long-standing feature of the active externalist literature,⁴ the nature of its relationship to cognitive extension remains unclear. Our aim in the present paper is to advance our understanding of this relationship. In Sect. 2, we survey the various appeals that have been made to transparency within the active externalist literature. In Sect. 3, we turn our attention to some of the problems that confront the effort to cast transparency as a criterion for cognitive extension. Finally, in Sect. 4, we present a theoretical account that seeks to explicate the role of transparency in arguments for extended cognition and the extended mind. Such an account, we suggest, helps us understand why phenomenal transparency might be deemed important for cognitive extension, even though it plays no role as a criterion for cognitive extension.

2 Cognitive extension and transparency

A number of theorists have suggested that transparency is relevant to cognitive extension (Clark, 2003, 2008a, 2015; Heersmink, 2015; Kirsh, 2019; Piredda & Di Francesco, 2020; Wheeler, 2019). Clark (2003), for example, makes a distinction between what he calls transparent and opaque technologies,⁵ with transparent technologies serving as the better candidates for cognitive incorporation.

In later work, Clark suggests that transparency is a subjective marker of cognitive incorporation, one that can be put to use in distinguishing genuine cases of cognitive extension from those of the more ersatz variety. The reason for this, Clark suggests:

[...] is because typical extended mind scenarios rely upon fluid unreflective use as one of the markers of incorporation into my own extended cognitive architecture, thus distinguishing true incorporation from mere—even if careful—tool use. (Clark, 2015, p. 3373)

Transparency also figures as part of Clark's response to one of the objections raised against the extended mind. This objection is what is commonly known as the Otto 2-Step (Clark, 2008a, p. 80). It concerns the claim that two individuals—Otto and Inga—can be said to possess the dispositional (or standing) belief that The Museum of Modern Art (MoMA) is located on 53rd Street, even though the storage and retrieval of belief-relevant information varies across the two individuals (while Otto relies on a trusty notebook to retrieve information about MoMA's location, Inga relies on her

⁴ For recent overviews of work in this area, see Andrada (2020, 2021), Wheeler (2019), and Facchin (2022).

⁵ According to Clark (2003, p. 37): "A transparent technology is a technology that is so well fitted to, and integrated with, our own lives, biological capacities, and projects as to become [...] almost invisible in use. An opaque technology, by contrast, is one that keeps tripping the user up, requires skills and capacities that do not come naturally to the biological organism, and thus remains the focus of attention even during routine problem-solving activity."

bio-memory). The Otto 2-Step challenges the idea that Otto ought to be credited with the dispositional belief that MoMA is on 53rd Street. According to the objection: “All Otto actually believes (in advance) is that the address is in the notebook. That’s the belief (step 1) that leads to the looking (step 2) that then leads to the (new) belief about the actual street address” (Clark, 2008a, p. 80). The worry, then, is that Otto does not possess the belief pertaining to the location of MoMA; rather, his belief pertains to the location of MoMA-related information.

In response to this worry, Clark makes an explicit appeal to transparency:

[...] to the worry that all that Otto really believes before consulting the notebook is that the address might perhaps be in the notebook, we reply that the notebook is consulted automatically and accessed without conscious deliberation, just like Inga’s biological recall. In this way, the notebook (like biological memory) functions as transparent equipment in a broadly speaking Heideggerian (1927/1961) sense. (Clark, 2005, pp. 2–3)

Here we can see how the notion of transparency is being invoked to address one of the criticisms that has been directed at the extended mind. In essence, the idea is that transparency helps to ensure that Otto is in possession of one sort of dispositional belief (i.e., the belief that MoMA is on 53rd Street) as opposed to another sort of dispositional belief (i.e., the belief that MoMA-related information can be found within the notebook). We can, of course, question the success of this particular response to the Otto 2-Step (see Wikforss, 2014); for present purposes, however, the point is that transparency seems to be playing a rather important role in arguments for the extended mind. In the absence of transparency, the door appears to be open to an alternative folk psychological characterization of the Otto notebook case, one that is largely inimical to claims of cognitive extension.

All this suggests that transparency might be of criterial relevance to cognitive extension. In respect of this issue, Wheeler (2019) suggests that we can regard transparency as either necessary or sufficient for cognitive extension, but he demurs from the idea that transparency is sufficient for cognitive extension on the grounds that no one subscribes to this view. What we are left with, then, is the idea that transparency is necessary for cognitive extension. Call this the *transparency constraint*. Here is how Wheeler presents the transparency constraint:

[...] given a situation in which an organic human being is using a tool, the transparency of that tool when in use is necessary for it to be a genuine part of that individual’s mental machinery—if the tool is not transparent, then what we confront is not a case of extended cognition. (Wheeler, 2019, p. 862)

Supporters of the transparency constraint include Thompson and Stapleton (2009). Their characterization of the constraint differs from that provided by Wheeler, but the general idea remains the same:

For anything external to the body’s boundary to count as a part of the cognitive system it must function transparently in the body’s sense-making interactions with the environment. (Thompson & Stapleton, 2009, p. 29)

Both these quotations express a commitment to the idea that transparency is necessary for cognitive extension. They also highlight a particular role for transparency—one that Wheeler (2019, p. 859) refers to as a “constituency condition.” The idea here is that transparency helps us determine when some bio-external resource forms a proper *part* (or constituent) of an individual’s cognitive system. In the absence of transparency, the bio-external resource amounts to nothing more than a mere tool with which an individual interacts.

Despite the various appeals to transparency within the active externalist literature, the relationship between transparency and cognitive extension remains obscure. As we have seen, some theorists have explicitly endorsed the idea that transparency is of criterial relevance to cognitive extension (e.g., Thompson & Stapleton, 2009). Other theorists, however, have suggested a somewhat weaker link between transparency and cognitive extension. Clark (2008a), for example, appears to distance himself from the transparency constraint when he writes that:

Experience is, of course, no more than a clue. I do not mean, here or elsewhere, to advance any arguments of the form ‘it seems to us as if we are/are not cognitively extended; therefore, we are/are not cognitively extended’! (Clark, 2008a, p. 238)

In this case, phenomenological considerations appear to be no more than an indication (a “clue”) as to the presence of cognitive extension. The problem with such proposals is that they leave us none the wiser as to the sort of role that transparency is playing in arguments for extended cognition and the extended mind. If transparency is relevant to cognitive extension, but this relevance is not to be understood as a form of criterial relevance (i.e., as a necessary or sufficient condition), then what sort of relevance is it?

We thus confront a problem regarding the link between transparency and cognitive extension. Should this link be construed as a form of criterial relevance, such that transparency is necessary for cognitive extension? If not, then how should we understand the appeal to transparency in the active externalist literature? What is motivating such appeals, and why might they be important?

In Sect. 4, we will hazard a response to these questions by outlining a particular approach to understanding cognitive extension. Before that, however, it is worth looking at some of the problems that confront the appeal to transparency in the active externalist literature. In the next section (Sect. 3), we review some of the reasons why transparency ought not to be seen as a (necessary) criterion for cognitive extension. We also raise doubts about the idea that transparency ought to be understood along the lines of a constituency condition; i.e., as something that enables us to determine when some bio-external resource ought to be seen as part of the machinery of the mind, or as a constituent of an individual’s cognitive system.

3 Problems with transparency

3.1 Mechanistic constitution

How ought we to understand the appeal to transparency in the active externalist literature? One possibility is that transparency is a distinctive feature of cognitive extension, one that is not encountered in cases of so-called embedded cognition (see Rupert, 2004). As noted by Wheeler (2019), the contrast between extended and embedded cognition is one that is typically understood with respect to constitutional and causal claims:

[...] in cases of extended cognition, the machinery of mind stretches beyond the skull and skin, in the sense that certain external elements are, like an individual's neurons, genuine constituents of the material realizers of that individual's cognitive states and processes [...] By contrast, in cases of what is now often called embedded cognition, the machinery of mind remains internal, but the performance of that inner mental machinery is causally scaffolded in significant ways by certain external factors. (Wheeler, 2019, p. 861)

It seems, then, that there is an important difference between extended and embedded cognition. Whereas the proponents of embedded cognition insist that all forms of cognitive extension can be understood with respect to the notion of causal relevance—the idea that bio-external resources exert a merely causal influence on cognitive states and processes—the proponents of extended cognition insist that there is something more than causal relevance at play in cases of cognitive extension. In particular, a distinctive feature of arguments for extended cognition and the extended mind is that bio-external resources can, at times, form a proper part of the material fabric that realizes cognitive states and processes. They are, in the words of Wheeler (2019, p. 857), “a constitutive part of our cognitive machinery.”

But what might it mean for a resource to count as a constitutive part of our cognitive machinery? One way of responding to this question is to turn our attention to work in so-called neo-mechanical philosophy, especially that which focuses on theories of mechanistic explanation (Craver, 2007b; Glennan, 2017; Glennan & Illari, 2018a). From a mechanistic standpoint, cognitive phenomena (e.g., cognitive states and processes) are explained by detailing the mechanisms that are responsible for those phenomena. What it means for a mechanism to be responsible for a phenomenon varies according to the type of (mechanistic) explanation that is provided for the phenomenon (see Kaiser & Krickel, 2017). In etiological (or causal) mechanistic explanations, the goal is to identify the mechanism that exerts a causal influence on the explanandum phenomenon (e.g., the mechanism that causes the phenomenon to occur). This contrasts with constitutive mechanistic explanations, where the goal is to identify the mechanism that constitutes, underlies, or realizes the explanandum phenomenon. In addition to this distinction between causal and constitutive explanation, there is a corresponding distinction between two forms of explanatory relevance (see Craver, 2007a). Etiological mechanistic explanations are thus concerned with issues of *causal relevance*—they seek to identify the objects and activities that exert a causal influence

on the to-be-explained phenomenon. Constitutive mechanistic explanations, by contrast, are concerned with issues of *constitutive relevance*. They seek to identify the objects and activities that serve as the constituents (or components) of a mechanism. It is by discerning such objects/activities, and charting their interactions with one another, that we come to understand the mechanistic basis of a phenomenon—how a phenomenon is realized by an interacting nexus of material objects.⁶

Inasmuch as we accept the idea that extended cognition ought to be understood from a mechanistic standpoint—i.e., as a claim about the constitutive relevance of bio-external resources to cognitive phenomena—then we have a relatively straightforward means of understanding the disagreement between extended and embedded theorists.⁷ Extended theorists insist that bio-external resources are, on occasion, constitutively relevant to cognitive phenomena and are thus *bona fide* constituents of the mechanisms that realize those phenomena. Embedded theorists, by contrast, insist that bio-external resources are merely causally relevant to cognitive phenomena—such resources may exert a causal influence on the mechanisms that realize cognitive phenomena, but they ought not to be seen as the constituents of such mechanisms. This dispute, it should be clear, is really one about the constitutive relevance of bio-external resources. If we were able to determine that bio-external resources were, indeed, constitutively relevant to cognitive phenomena, then the debate would be settled in favor of the active externalist camp (see Kaplan, 2012).

To the extent that phenomenal transparency has any bearing on this debate, it must be because transparency is relevant to the distinction between constitutive and causal relevance. Accordingly, the transparency constraint might be cast as a criterion of constitutive relevance: In order for a bio-external resource to count as a *bona fide* constituent of a cognitive mechanism, it must meet the conditions of phenomenal transparency; that is to say, the resource must be transparent-in-use. If this is not the case, then the resource is not so much a constituent of the mechanism as it is a causally-relevant factor that influences the operation of that mechanism.

Unfortunately, there are reasons to think that this approach to transparency cannot be correct. The problem relates to the generality of constitutive relevance—the fact that issues of constitutive relevance arise in respect of *all* mechanisms, not just those of the (extended) cognitive variety. The causal/constitutive distinction is thus something that applies to explanatory contexts where the appeal to transparency makes little or no sense. Consider, for example, that the notion of constitutive relevance is just as

⁶ See Craver (2007a), for more on the distinction between constitutive and causal relevance. For a discussion of constitutive/causal relevance in an active externalist context, see Kaplan (2012). For a recent philosophical account of constitutive relevance, see Craver et al. (2021).

⁷ This assumes, of course, that mechanistic concepts are applicable to our understanding of extended cognition and the extended mind. Support for this idea stems from the terminology used by the proponents of active externalism—consider, for example, Wheeler's (2019) reference to cognitive/mental machinery. It also stems from the emphasis assigned to explanatory practices as a means of understanding cognitive extension (Fazekas, 2013; Hurley, 2010). Finally, and perhaps most importantly, there is a general sense that the notion of extended cognition is best understood with respect to what have been dubbed *extended* (Clark, 2011; Hurley, 2010; Kaplan, 2012; Smart, 2022; Zednik, 2011), *wide* (Milkowski et al., 2018) or *supersized* (Clark, 2008a) mechanisms. The idea here is that a cognitive process ought to be seen as extended if it is realized by an extended (wide or supersized) mechanism, where an extended mechanism is one that includes resources that lie external to the cranial/corporeal boundary.

applicable to the mechanistic explanation of social and astrophysical phenomena as it is to the mechanistic explanation of cognitive phenomena (Illari & Williamson, 2012; Ylikoski, 2018). This presents us with a significant problem: Inasmuch as we are able to resolve issues of constitutive relevance in respect of non-cognitive phenomena, then it is hard to see why transparency ought to be seen as a prerequisite for constitutional claims. It would make no sense, for example, to insist on transparency in the case of astrophysical mechanisms. Similarly, it would make no sense to insist on transparency in the case of mechanisms with so-called designed-and-built etiologies (Glennan & Illari, 2018a). If we can determine that a piston forms part of the propulsive mechanism for a conventional automobile, and we can do so without an appeal to transparency, then it must be possible for us to resolve issues of constitutive relevance without invoking the notion of transparency. But, if that is the case, then it is hard to see why transparency ought to be understood as something like a “constituency condition” for cognitive extension (see Wheeler, 2019, p. 859). It is possible that the notion of a constituency condition ought to be understood in some other way; i.e., in a way that does not appeal to constitutive relevance. But given the widespread appeal to mechanisms, machinery, and the causal/constitutive framing of the debate between extended and embedded theorists (see above), it is difficult to know what this alternative interpretation might be.

3.2 Exotic forms of cognitive extension

For the most part, debates about extended cognition and the extended mind have been oriented to the realm of human cognizing. In particular, the majority of debates and discussions limit their focus to situations where human individuals are interacting with one or more bio-external resources. These are what we might call *human-centered forms of cognitive extension*. They are human-centered in the sense that the subject of extension—the entity whose cognitive states and processes are deemed to be extended—is the biologically-bounded human individual.

Human-centered forms of cognitive extension are not, however, the only forms of cognitive extension that have been discussed in the philosophical and cognitive scientific literature. Other forms of cognitive extension include those centered on plants (Parise et al., 2020), spiders (Japyassú & Laland, 2017; Smart et al., 2010), slime moulds (Sims & Kiverstein, 2022), and non-biological Artificial Intelligence (AI) systems (Smart, 2018). It should be clear that these forms of cognitive extension present a significant problem for those who deem transparency to be necessary for cognitive extension. It is, in particular, difficult to see how we might apply the notion of transparency to something like a spider—an entity whose phenomenology (if it has any) is inaccessible to us.

As with the foregoing discussion of constitutive relevance, the presence of exotic (or, at any rate, non-human) forms of cognitive extension challenges the idea that transparency ought to be seen as necessary for cognitive extension. Insofar as we allow for the possibility of non-human forms of cognitive extension, then it seems that some species of extended cognizing can be discerned without the appeal to transparency. But, if that is the case, then it is hard to see why we ought to accept the transparency

constraint. If the goal is to determine whether some extra-organismic resource forms part of the cognitive machinery of some entity, and we can resolve this issue without the appeal to transparency in the case of non-human entities, then what is the basis for insisting on transparency in those situations where the subject of extension just so happens to be a human individual?

We could, of course, reject the possibility of non-human forms of cognitive extension. Alternatively, we could try to alter the scope of the transparency constraint, limiting it to the realm of human (personal-level⁸) cognizing. To our mind, neither of these strategies are likely to work. While we might want to contest the claims that have been made about extended cognition in spiders, plants, and other entities; we surely do not want to reject the mere possibility of non-human forms of cognitive extension. As regards the second strategy, it is unclear why we would want to limit the transparency constraint to the specific realm of *human* cognizing. (Presumably, this cannot have anything to do with the fact that the transparency constraint can only be evaluated in human contexts.) To our mind, both these strategies are apt to raise concerns about the role of certain types of human experience in defining the scope of extended cognitive science. Active externalism has done much to challenge the way we think about the mind, encouraging us to abandon neurocentric and bio-chauvinistic prejudices in favor of a more neutral perspective. Such contributions are both important and welcome, but it would be unfortunate if the price of one form of neutrality were to be another form of prejudice. If all we are doing is abandoning neurocentrism in exchange for a highly dubious (and species-specific) form of experiential chauvinism, then perhaps not so much has been gained.

Suppose, however, that we were to limit the notion of cognitive extension to the realm of human (personal-level) cognizing. Would this actually help to address the sort of worries raised by non-human forms of cognitive extension? To our mind, the answer is “no.” To help us see this, consider the classic extended mind case involving Otto and the notebook. On the one hand, we can certainly entertain assumptions about Otto’s phenomenology, including the way Otto experiences his notebook. But, then again, none of us really knows what it is like to be Otto. And what difference does it make, in any case? Suppose that Otto is a philosophical zombie (see Kirk, 2015), devoid of any sort of phenomenology. Does this have any bearing on the extent to which we can ascribe dispositional beliefs to Otto as a means of establishing a folk psychological grip over his overt behavior? Likewise, what difference would it make if Otto’s interactions with his notebook were *not* to be accompanied by the phenomenal experience of transparency? Providing this makes no difference to Otto’s overt behavior, is there a reason why our folk psychological glosses would be rendered invalid in this situation? As noted in Sect. 2, transparency features as part of Clark’s response to the Otto 2-Step. But the origins of this problem have more to do with the granularity of folk psychological explanations⁹ and the conditions under which it makes both social and subjective sense to credit individuals with certain (dispositional)

⁸ Note that a number of sub-personal forms of cognitive extension have been discussed in the literature (Boem et al., 2021; Facchin et al., 2021). While we do not discuss these in the present paper, such forms of cognitive extension are also apt to raise problems for the proponent of the transparency constraint.

⁹ See Chalmers’ discussion of this issue in the foreword to Clark (2008a, pp. xii–xiii).

beliefs. Transparency is arguably important, here, especially when it comes to the way we recognize ourselves as being in possession of certain abilities, beliefs, and knowledge. But the overarching concern, here, is not so much transparency *per se*; it is more that certain properties (beliefs, abilities, knowledge, etc.) are being ascribed to particular individuals. Inasmuch as such properties can be ascribed in the absence of transparency (or the absence of any sort of insight into phenomenological properties), then the appeal to transparency looks to be beside the point.

3.3 Phenomenal parity and commonsense functionalism

In recent years, there has been a growing interest in the epistemological implications of active externalism. In particular, there has been a growing interest in the possibility of extended knowledge—the idea that knowledge can stem from the operation of extended cognitive processes, or that epistemic states can be extended in precisely the same way as states of dispositional belief (see Carter et al., 2018a). Interestingly, this debate has sparked a renewed interest in transparency-related issues. According to some theorists, extended knowledge is tied to issues of awareness, such that extended knowers need to be aware of the reliability of their (extended) belief-forming processes (see Pritchard, 2010). The problem with such proposals is that the epistemically-motivated appeal to awareness threatens to undermine claims of cognitive extension. With respect to the Otto notebook case, for example, Clark (2015) suggests that:

[...] the more he [Otto] is aware of such matters [e.g., the reliability of the belief-forming process], the less the notebook will seem to be playing the same kind of functional role as biological memory. For [...] our biological memory is not typically subject to agentive scrutiny as a process at all, much less as one that may or may not be reasonably judged to be reliable by the agent. (Clark, 2015, p. 3763)

What we see here is an appeal to phenomenological criteria in evaluating putative cases of cognitive extension. In particular, it seems as though the phenomenology of the extended mind should approximate that of the non-extended mind, such that if a bio-external resource (e.g., a notebook) is to count as a constituent of an extended mind, then it ought to be experienced in much the same way as we experience our own biologically-based cognitive machinery (e.g., our bio-memory system). The notion of “experience” is, of course, somewhat broader than the notion of transparency; nevertheless, transparency does appear to be one of the factors that figures in this call for phenomenal parity. Clark (2003), for example, suggests that:

Transparent technologies are those tools that become so well fitted to, and integrated with, our own lives and projects that they are [...] pretty much invisible-in-use. These tools or resources are usually no more the object of our conscious thought and reason than is the pen with which we write, the hand that holds it while writing, or the various neural subsystems that form the grip and guide the fingers. All three items, the pen, the hand, and the unconsciously operating neural mechanisms, are pretty much on a par. *And it is this parity*

that ultimately blurs the line between the intelligent system and its best tools for thought and action. (Clark, 2003, pp. 28–29; emphasis added)

In one sense, there is nothing problematic about any of this: we are simply relying on phenomenological criteria to evaluate putative cases of cognitive extension. There is, however, a significant problem with the notion of phenomenal parity. The problem relates to the functionalist flavor of arguments for the extended mind. In particular, arguments for the extended mind are based on the idea that bio-external resources should function in a manner that is compatible with the folk psychological apparatus of thought ascription. What this means, in effect, is that a bio-external resource (e.g., a notebook) should influence thought and action in the manner we typically expect of a folk psychological state, such as a state of dispositional belief. If it does this, then the resource's functional role will approximate that of a purely brain-based resource such as a bio-memory system, and it is this comparison with bio-memory that is driving arguments for the extended mind: Given that we are perfectly content to accept bio-memory as part of an individual's cognitive system, then we ought to do the same for a functionally-equivalent bio-external resource. Accordingly, we ought to judge putative cases of cognitive extension (specifically, the extended mind) based on the functional parity of the inner and outer resources.

The problem is that it is not entirely clear how this commitment to functional parity relates to the earlier commitment to phenomenal parity. If what matters in extended mind cases is simply the idea that some bio-external resource is influencing thought and action in the manner we typically expect of a folk psychological kind (i.e., a state of dispositional belief), then it is hard to see why we ought to be concerned about the phenomenological peculiarities of whatever processes are associated with the retrieval of belief-relevant information. As noted by Andrada (2021), it is far from clear that phenomenal differences will always translate into functional differences, such that all forms of phenomenal divergence will have a bearing on issues of functional parity. Providing the notebook entries provide us with a folk psychological grip over Otto's actual and counterfactual behavior, then the notebook will count as a *bona fide* constituent of Otto's extended mind. In this case, the presence/absence of transparency looks to be of little consequence. To echo one of the points mentioned above (see Sect. 3.2), we can ask ourselves whether our intuitions about the Otto case would be any different if Otto's behavior should remain the same despite the absence of transparency? It is hard to see why this should be so, especially since we are being asked to disregard or paper over any number of other differences between the two protagonists of the Otto case (i.e., Otto and Inga).¹⁰

Let us return to Clark's worry about the appeal to awareness in the context of claims about extended knowledge. Clark's worry, recall, is that the "more [Otto] is aware of such matters, the less the notebook will seem to be playing the same kind of functional role as biological memory" (Clark, 2015, p. 3763). What Clark seems to be suggesting here is that we ought to see phenomenal parity as a component of functional parity. That is to say, what it means for the notebook to fulfill the functional role of a bio-memory

¹⁰ In respect of this issue, Clark (2011, p. 451) suggests that we ought to ignore a number of experimentally-documented features of bio-memory, such as primacy and recency effects. Such effects, Clark suggests, are irrelevant to the sort of functional role that motivates claims about the extended mind.

system is that it is encountered as a form of transparent equipment. This, however, is a highly contentious claim. It is, in particular, unclear whether phenomenological talk is really compatible with the notion of a functional role. Is it appropriate, for example, to suggest that part of the functional role of a memory system is that it be transparent-in-use? To our mind, this seems to stretch the notion of a functional role beyond the limits of what is suggested by extended functionalism—the sort of functionalism that is favored by proponents of the extended mind (Clark, 2008b; Wheeler, 2010).

There are other reasons to be skeptical about the appeal to phenomenal parity. Andrada (2021), for example, questions the extent to which our (folk-theoretic) understanding of memory is wedded to the notion of phenomenal transparency. In respect of this issue, Andrada (2021) suggests that the core functional features of a memory system are that it supports the selective storage and context-sensitive retrieval of information. This, she suggests, is perfectly compatible with a number of agential attitudes, including ones where the agent is aware of the memory system and the reliability of the information contained therein.

It is also worth considering the way that phenomenological issues have been discussed in situations where epistemic concerns (e.g., the appeal to extended knowledge) are not particularly paramount. Consider, for example, the way that Clark and Chalmers (1998) discuss one point of phenomenological divergence between Otto and Inga:

[In Otto's case] there is a distinctly perceptual phenomenology associated with the retrieval of the information, whereas in Inga's case there is not. But why should the nature of an associated phenomenology make a difference to the status of a belief? (Clark & Chalmers, 1998, p. 16)

The question that arises here is why phenomenological differences should matter in the case of extended knowledge if we are already dismissing (or at least downplaying) such differences in the case of extended belief? Clark (2015) suggests that the appeal to agential attitudes, active scrutiny, conscious apprehension, and so on, creates a point of tension in the debate over extended knowledge. To our mind, however, there is another sort of tension at play here. It concerns the apparent inconsistency in claims regarding the putative importance of phenomenological factors in arguments for the extended mind. If it is indeed the case that perceptual phenomenology has no bearing on arguments for the extended mind, then what is the basis for claiming that phenomenal transparency ought to be treated any differently?

3.4 Now you see it; now you don't

According to Wheeler (2019):

Taking transparency to be necessary for cognitive extension underlies a dynamic version of ExM [the Extended Mind Hypothesis]. Minds grow beyond the skin and shrink back to the boundary of the skin, depending, in part, on the phenomenological dynamics of our couplings with technology. When a tool is transparent, that is a necessary condition met for its constitutive incorporation

into the user's mental machinery. When a tool becomes visible, due to, for example, damage or malfunction, or when, as in the case of some sensory substitution subjects, a deliberate, conscious effort on the part of the user resets the mind-world boundary at the skin, that means that cognitive extension is no longer operative. (Wheeler, 2019, p. 862)

As Wheeler notes, if transparency is necessary for cognitive extension then a temporary shift in transparency threatens to reset the borders of the mental machine, such that an erstwhile constituent of one's cognitive system suddenly assumes the status of an external (non-incorporated) object. For the sake of convenience, let us dub this the *shifting boundaries thesis*. What is crucial to this thesis is not so much the idea that the borders of a cognitive system are apt to expand and contract in certain circumstances (for this is something that is already entailed by the notion of cognitive extension). What is distinctive about the shifting boundaries thesis is that the borders of a cognitive system are apt to vary according to the presence or absence of phenomenal transparency—or, as Wheeler puts it, minds grow and shrink according to the “phenomenological dynamics of our couplings with technology.”

As noted by Wheeler, the shifting boundaries thesis follows from the transparency constraint. Thus, if one subscribes to the view that transparency is necessary for cognitive extension, then a shift in transparency will reconfigure the borders of the mental/cognitive machine. Wheeler does not seem to regard this as particularly problematic: We should, he suggests, “be untroubled by the idea that each of us possesses a dynamically growing and shrinking extended mind” (Wheeler, 2019, p. 862). From our perspective, however, there are reasons to doubt the tenability of the shifting boundaries thesis, and this leads to a possible *reductio* of the idea that transparency is necessary for cognitive extension.

One reason to reject the shifting boundaries thesis stems from the purported link between transparency and “constitutive incorporation”—the idea that transparency is something that influences the extent to which some resource ought to be seen as part of an individual's cognitive/mental machinery. We have already subjected this idea to critical scrutiny in earlier sections. Thus, in Sect. 3.1, we suggested that transparency cannot serve as a criterion of constitutive relevance, and this raises doubts about its status as a constituency condition. Does this mean that we are in a position to reject the shifting boundaries thesis?

The answer to this question is complicated by the fact that there is a way of understanding the shifting boundaries thesis that is perfectly compatible with the notion of constitutive relevance. This compatibility centers on the idea that one of the things that can disrupt transparency is the presence of some sort of damage or malfunction. To help us understand this, suppose you are using pen and paper resources to solve a long multiplication problem.¹¹ While you are actively engaged in the problem-solving effort, the pen and paper resources are not at the forefront of your conscious awareness, and they thus qualify as transparent equipment. This will, of course, change if something should go awry. If the pen should break, for example, your attention will be drawn to it. In this situation, the pen no longer counts as transparent equipment, and

¹¹ The use of pen and paper resources to solve long multiplication problems is a frequently cited example of extended cognizing [e.g., Wilson and Clark (2009) and Wheeler (2010)].

thus the pen no longer counts as a constituent of the multiplicative process, at least according to the shifting boundaries thesis. As it turns out, this conclusion is correct, but the reason it is correct has nothing to do with transparency. It is more that the breaking of the pen disrupts the performance of the long multiplication routine and thus the pen no longer qualifies as a constituent of the mechanism that realizes the routine. While the pen is broken or lying idly on the desk, it is not part of a mechanism that realizes a token instantiation of the long multiplication process; it only becomes a constituent of such a mechanism when it is picked up and pressed into useful mathematical service. To be sure, then, a shift in transparency may accompany a change in the borders and boundaries of a cognitive/mental machine, as when the performance of an ongoing cognitive routine is interrupted due to damage or malfunction. But we can understand this shift in mechanistic borders and boundaries without appealing to the notion of transparency. The mistake here is to see a shift in transparency as contributing to the dissolution of an extended cognitive organization, when, in fact, it is the (perhaps unexpected) dissolution of that organization that is prompting a shift in transparency.

As noted in Sect. 3.1, it is possible that the notion of a constituency condition ought to be understood in a way that does not appeal to constitutive relevance. Perhaps, for example, the sort of constituency that Wheeler has in mind is better understood with respect to the sorts of things that make us the particular cognitive agents we are. Something would thus count as part of me (and thus part of my cognitive system) if it played a role in supporting the various cognitive/mental properties (e.g., cognitive abilities, beliefs, and knowledge) that are the features of my own cognitive/mental character.¹² Perhaps, then, we can make sense of the link between transparency and constituency via the sort of ‘parthood’ that is at play when it comes to subjective assessments of our own abilities, beliefs, and knowledge. If I deem myself able to solve long multiplication problems, and the exercise of this ability is one that relies on the use of bio-external resources (e.g., pen and paper artifacts), then the bio-external resources are clearly relevant to the sort of ability that I take myself to have. Divested of those resources, I might well be bereft of *my* multiplicative capacities, and thus the bio-external resources count as ‘part’ of me because they are part of the material fabric that sustains my subjective sense of the sort of cognitive entity I am.

We will have more to say about this in Sect. 4; for present purposes, however, it is worth bearing in mind that this alternative approach to constituency does not offer much in the way of support for the shifting boundaries thesis. The primary problem is that our sense of who and what we are appears to be resistant to temporary shifts in transparency. Given the reliable presence of pen and paper resources, I may come to regard myself as someone who is able to solve long multiplication problems. But this sense of who I am (and what I can do) is not particularly affected by the sort of events that are apt to prompt a shift in phenomenal transparency during the performance of

¹² This is the sort of constituency that has surfaced in recent debates about the extended self. Clowes (2020), for example, suggests that our sense of who and what we are is informed by the sorts of skills and capacities that we take ourselves to have, but these skills and capacities may be ones that are tied to the reliable presence of a rich array of bio-external resources. From this perspective, then, a bio-external resource would count as ‘part’ of an individual’s cognitive system if it played a role in sustaining the cognitive skills and capacities of that individual.

a cognitive task. If the pen should break while I am engaged in a long multiplication routine, then I can just get another. My sense of who I am, what I can do, and what I know remains largely unaffected by these short-term shifts in my phenomenological orientation to bio-external resources. If this is the sort of way we are being asked to think about transparency (i.e., as something that informs our sense of who and what we are as cognitive agents), then it seems that the shifting boundaries thesis cannot be correct. And since the shifting boundaries thesis follows directly from the transparency constraint, then the validity of the transparency constraint is itself called into question.

4 Understanding transparency: a working hypothesis

The previous section highlighted a number of problems for the idea that transparency is necessary for cognitive extension. In the wake of these problems, we are inclined to reject the transparency constraint. Transparency, we suggest, should not be seen as necessary for cognitive extension.

A further problem relates to the idea that transparency can be used to resolve disputes about the constitutive (versus merely causal) status of bio-external resources (see Wheeler, 2019). Relative to the discussion in Sect. 3 (and, especially, Sect. 3.1), we see little in the way of support for this idea. Inasmuch as we are to understand cognitive extension from a broadly mechanistic perspective, then transparency cannot be used as a criterion of constitutive relevance.

All this leave us with a problem: If transparency is not to be understood in criterial terms (i.e., as a criterion for cognitive extension), and we also reject the idea that transparency is relevant to constitutional matters, then how are we to understand the appeal to transparency in the active externalist literature?

Our response to this challenge comes in the form of what we will call the *dispositional hypothesis*. According to this hypothesis, claims about extended cognition and the extended mind can be understood in the following way: In cases of cognitive extension, we ascribe a dispositional property¹³ (D) to an entity (E), but the manifestation/exercise of D is a phenomenon (P) that is subject to extended mechanistic realization.¹⁴ What we mean by extended mechanistic realization is that the mechanism (M) responsible for P includes components that lie beyond the borders/boundaries (B) of E (i.e., beyond the borders/boundaries of the thing to which a dispositional property is ascribed). Applying this to the traditional target of active externalist theorizing, namely, human-centered forms of cognitive extension, we arrive at the following:

¹³ Note that we are adopting a property-based approach to dispositions. That is to say, we are claiming that dispositions are properties. This is consistent with the philosophical analysis of dispositions in a scientific context [e.g., Hüttemann and Kaiser (2018)]. For an alternative approach to dispositions, see Mumford (2009).

¹⁴ In particular, the phenomenon (P) is a constitutive mechanistic phenomenon: a process, event, state, or other occurrent that is exhibited by a given entity, object, or system [see Kaiser and Krickel (2017)].

- E: a human individual;
- D: a cognitive/mental dispositional property (e.g., a cognitive ability, cognitive capacity, or dispositional belief);¹⁵
- P: a cognitive process or mental state; and
- B: the biological (skin/skull) boundary of the human individual.

The dispositional hypothesis is sufficiently generic to accommodate many of the cases that have been discussed in the philosophical literature. In the case of the extended mind, for example, D is a dispositional belief, and P refers to the process of retrieving information from a bio-external resource (e.g., a notebook). In other situations, D corresponds to a cognitive ability, such as an ability to solve long multiplication problems using pen and paper resources. In this case, we have a form of extended cognizing due to the fact that the pen and paper resources are the constituents of a mechanism (M) that is responsible for the realization of a multiplicative routine (P), where the multiplicative routine corresponds to the manifestation (or exercise) of a cognitive ability (D) that is ascribed to a particular human individual (E).

For reasons of space, we will not attempt to discuss the details of the dispositional hypothesis in the present paper; our primary aim is to assess the extent to which this hypothesis might be used to illuminate the link between phenomenal transparency and cognitive extension.¹⁶ With this in mind, let us contrast two situations in which a human individual is using a technological device to perform a cognitive task. In one case (call it the *fluent case*), the user has some prior experience with the technological device and uses it in a fluent and effortless manner. In the other case (call it the *non-fluent case*), the user has little or no prior experience with the device and struggles to use it. Proficiency is thus one of the things that varies across the two cases. But given that proficiency has been associated with transparency, let us also assume that the two cases differ with respect to the presence of phenomenal transparency. Accordingly, let us assume that the technological device counts as transparent equipment in the fluent case, while, in the non-fluent case, it fails to count as transparent equipment.

Inasmuch as we want to claim that the fluent case ought to be seen as a *bona fide* form of extended cognizing (while the non-fluent case is not), then we will obviously need some way of distinguishing between the two cases. Clearly, the appeal to phenomenological factors—the appeal to transparent equipment—is one way of making this distinction. But it is not the only way that such a distinction might be made.

¹⁵ For present purposes, we will assume that abilities, capacities, and dispositional beliefs are particular kinds of dispositional property. Other kinds of dispositional property include tendencies, proclivities, capabilities, propensities, and potentialities.

¹⁶ The dispositional hypothesis is, we suggest, broadly consistent with the way that cognitive extension has been discussed in the philosophical literature. Wilson and Clark (2009), for example, draw attention to the importance of dispositional properties (specifically, cognitive capacities) in their discussion of so-called coupling conditions. The purpose of these conditions, they suggest, is to support the ascription of cognitive capacities to a given individual, even though the capacities in question are ones that rely on a distributed nexus of forces and factors, some of which lie external to the individual's biological boundary: "Such coupling conditions are meant to ensure that the capacities of the hybrid system—the biological organism plus augmentation—are plausibly seen as the capacities of a specific individual (e.g. Otto). We properly expect our individual agents to be mobile, more or less reliable, bundles of stored knowledge and computational, emotional, and inferential capacities. So we need to be persuaded that the new capacities enabled by the addition of the notebook are likewise sufficiently robust and enduring as to contribute to the persisting cognitive profile that we identify as Otto the agent" (Wilson & Clark, 2009, p. 67).

Our proposal is that the fluent and non-fluent cases are differentially effective in eliciting the sort of disposition ascriptions that (according to the dispositional hypothesis) serve as the basis for claims of cognitive extension. In the fluent case, for example, it seems perfectly appropriate to regard the human individual as possessing certain abilities. Clearly, the user has an ability to interact with the technological device for the purpose of producing certain outcomes. But we might be inclined to go beyond this and credit the individual with another sort of ability—an ability (or perhaps a capacity) to achieve certain sorts of cognitive outcome. (Consider, for example, that I take myself to possess an ability to solve long multiplication problems, and this is so even though the ability in question is one that relies on my facility with a range of bio-external resources.) Things are much less clear in the non-fluent case. In this case, our ability-based ascriptions are hampered by the fact that the user is struggling to use the technological device. The fluent and non-fluent cases thus vary with regard to their phenomenological characteristics (transparency vs. non-transparency), but they *also* vary with regard to the ascription of certain sorts of dispositional properties (in this case, abilities). This is important, for it is, we suggest, this latter difference that holds the key to understanding the appeal to transparency in the active externalist literature. In short, we suggest that transparency is a phenomenological feature of situations in which we ascribe certain cognitive/mental dispositional properties to both ourselves and others. From this perspective, transparency has nothing to do with constitutional matters (or it is, at any rate, only indirectly related to such matters). Rather than being relevant to constitutional matters, transparency is a phenomenological property of situations in which we credit ourselves and others with the possession of certain dispositional properties—properties that, when manifest, are realized by mechanisms that extend beyond the borders of the individual (or other entity) to which those properties are ascribed.

It is important to note that there are two kinds of disposition ascription that might be applied to the individual in the fluent case. The first of these is the ability to interact with the technological device for the purpose of completing a cognitive task. As noted above, however, there is another sort of ability that might be ascribed to the human user, namely, an ability to produce, accomplish, or achieve a certain outcome using the technological device. (Both these abilities, it should be clear, are absent in the non-fluent case.) The difference between these ability-based ascriptions is admittedly subtle, but it is nevertheless important. Consider that there is an important difference between an ability to utilize pen and paper resources for the purpose of solving long multiplication problems versus a ‘simple’ ability to solve long multiplication problems. The two are, of course, related: in the absence of the former, it is hard to see how we could be said to possess the latter. Nevertheless, when it comes to claims about cognitive extension, it is, we suggest, the latter ability that is the important one. The reason for this is that the latter ability plausibly counts as an *extended* ability. That is to say, it is an ability-based dispositional property (D) that is ascribed to a human individual (E), but the exercise of this ability (the disposition manifestation) is a process (P) (the long multiplication process) that is realized by a mechanism (M) that extends

beyond the borders of E.¹⁷ This differs from the former (non-extended) ability; i.e., the ability to interact with a given bio-external resource. In this case, the exercise of the ability is one that relies on mechanisms that are internal to the individual's organismic boundary; i.e., the neural and skeletomuscular mechanisms that support interaction with a bio-external resource. Another way of thinking about this is to recognize that there are two kinds of explanatory target associated with the fluent case: the first relates to the question of how an individual is able to interact with a bio-external resource (e.g., a technological device), while the second relates to the question of how an individual is able to (successfully) produce a certain outcome. These questions are apt to yield different answers as regards the spatial extent of the mechanisms that are responsible (in a constitutive sense) for the exercise or manifestation of the corresponding abilities that are ascribed to the human individual.

Now note something important: the details of our interactions with a bio-external resource are likely to be particularly salient in situations where phenomenal transparency is either absent or interrupted. Consider that if you can see me struggling to use a technological device, then you will be much less inclined to say that I have an ability to produce (or achieve) whatever it is that stems from the proficient use of that device. The issue here is not so much whether I have an ability to produce a certain outcome; it is more the issue of whether I possess an ability to interact with a bio-external resource in an appropriate manner. Similarly, if I can 'see' myself struggling to use a technological device, then my attention will be drawn to the details of my interaction with the device. In such a situation, it becomes difficult to credit myself with a simple ability to produce (or achieve) whatever outcome stems from the proficient use of the device, for it is precisely this proficiency that is called into question by *my* apparent struggling. Contrast this with a state-of-affairs in which the technological device is transparent equipment for me. In this case, the details of my interaction with the device are not the focus of my attention/awareness, and there is, as such, no need for me to reflect on my ability to use the device (my possession of this ability is, in fact, entailed by the transparent nature of my exchanges with the device). Instead of focusing on one sort of dispositional property (an ability to use a resource), I 'see through' the details of my interaction with the device to another sort of ability, namely, an ability to produce or achieve a certain outcome.¹⁸ Proficiency with bio-external resources thus helps us overlook one sort of ability (an ability to use a resource for the purpose of producing an outcome) in favor of another sort of ability (an ability to produce an outcome using a resource).¹⁹ And is this latter sort of ability, we suggest, that forms the basis for

¹⁷ This does not mean that *all* the components of M are external to E. In the long multiplication case, for example, it is likely that E will count as one of the components of M.

¹⁸ This does not mean that transparency is something that applies to abilities (and other dispositional properties). In short, we are not suggesting that abilities, themselves, are either transparent or non-transparent. It is more that phenomenal transparency gives us no reason to reflect on the more basic suite of (biologically-based) abilities that enable us to press bio-external resources into useful cognitive service. Thus, instead of saying that our mathematical successes are due to our facility with pen and paper resources, we simply say that we are in possession of certain mathematical abilities—rather than say we have an ability to use pen and paper resources to solve long multiplication problems, we simply say that we are able to solve long multiplication problems.

¹⁹ There is an interesting parallel here with the Otto 2-Step. The Otto 2-Step, recall, centers on the tension between the ascription of two sorts of dispositional property: (1) the belief that museum-related information

claims about extended cognizing. The phenomenological property of transparency is typically associated with situations where our interaction with a bio-external resource is glossed as expert, skilled, proficient, fluent, and trouble-free, but these are precisely the sorts of situations where we credit ourselves and others with the possession of dispositional properties that are (according to the dispositional hypothesis) subject to extended mechanistic realization.

This approach to transparency provides us with a means of making sense of the ostensible distinction between mere tools and genuine mind parts—a distinction that is particularly prominent in debates about extended knowledge (Carter et al., 2018b; Clark, 2015). We are thus happy to accept that the aforementioned fluent case is a candidate case of cognitive extension, whereas the non-fluent case is not. Crucially, however, this distinction has nothing to do with phenomenal transparency per se; it is more to do with the fact that the ascription of certain dispositional properties relies on the skilled, proficient, and expert use of bio-external resources. The more difficult it is for a cognitive agent to exploit an external resource, the less inclined we are to ascribe cognitive abilities/capacities to that entity, and without these abilities/capacities there is no *extended* cognizing. There is no extended cognizing because there is no cognitive ability/capacity that might be subject to extended mechanistic realization. That is to say, there is no ability/capacity that, when manifest, is realized by the coordinated flow of energy and information across a causally-interacting nexus of material objects, some of which lie beyond the borders of the thing to which the ability/capacity is ascribed.

One of the virtues of the dispositional hypothesis is that it provides a common approach to understanding extended cognition and the extended mind. In fact, all that really changes in the shift from one form of cognitive extension to the other is the nature of the dispositional property.²⁰ Thus, in the Otto notebook case, our attention shifts from the realm of cognitive abilities to the realm of dispositional beliefs, but this shift has no real bearing on the way we understand the case. Nor does this shift in the nature of the dispositional property radically alter the way we understand the appeal to transparency. Just as we need to be sure that it makes sense to credit an individual with a given cognitive ability/capacity, we also need to be sure that it makes sense to credit Otto with certain dispositional beliefs, specifically, those that are ‘contained’ in his notebook device. What this means, in effect, is that Otto’s thoughts and actions need to be coordinated with the notebook’s contents in such a way as to legitimate the ascription of certain (dispositional) beliefs to Otto. It would make no sense to say that Otto believes that MoMA is on 53rd Street if, when push comes to shove, Otto does not behave in a manner that is consistent with the ascription of this belief. If we ask Otto about MoMA’s location, then he should respond by saying “53rd Street.” Similarly, if Otto desires to go to MoMA, then he should respond by going to 53rd

Footnote 19 continued

can be found in the notebook, and (2) the belief that MoMA is on 53rd Street. Clark (2008a) responds to this problem by invoking the notion of transparency as a means of deflecting attention away from the former (non-extended) belief towards the latter (extended belief). This resembles our own approach to transparency, in the sense that transparency is helping us overlook one sort of disposition ascription in favor of another.

²⁰ As noted in Sect. 1, claims about extended cognition are typically formulated with respect to explanatory kinds that are relevant to cognitive science (e.g., cognitive abilities and cognitive capacities). This contrasts with claims about the extended mind, which are typically formulated with respect to explanatory kinds that are relevant to folk psychology (e.g., states of dispositional belief).

Street. If this should not be the case—if Otto should desire to go to MoMA, but then end up going to 43rd Street—then we would have little reason to credit Otto with the belief that MoMA is on 53rd Street. At the very least, there would be little reason to regard the notebook entries as providing us with any sort of folk psychological grip over Otto's behavior.

At a minimum, then, the contents of the notebook must serve as a reliable guide to Otto's actual and counterfactual behavior. This, we suggest, helps us understand the appeal to automatic endorsement in arguments for the extended mind (Clark & Chalmers, 1998; Clark, 2010). Automatic endorsement is important, for we want to avoid situations in which Otto is inclined to question the notebook entries, as might be the case if Otto deems the notebook to be an unreliable source of museum-related information. If Otto should deem the notebook to be unreliable, then he may question (and quite possibly reject) the information that is contained therein. This, however, presents us with a significant problem, for it now becomes unclear what Otto will do in those situations where the possession of a dispositional belief is evidenced by his overt behavior. If, for example, Otto should read that MoMA is on 53rd Street, but then decide that the information cannot be correct and go somewhere other than 53rd Street, then there would be nothing to substantiate the claim that Otto believes (in either a dispositional or occurrent sense) that MoMA is on 53rd Street. It is precisely for this reason, we suggest, that transparency looks to be important for the extended mind. For it is by insisting on transparency that we reduce the risk of Otto subjecting the notebook entries to critical scrutiny. Such forms of scrutiny (e.g., wondering whether or not the notebook entries are correct) open the door to situations in which Otto's thoughts and actions might run counter to what is written in the notebook. But if that should be the case, then the notebook entries will fail to provide us with any sort of folk psychological grip over Otto's actual and counterfactual behavior. The problem, here, it should be clear, is not so much the extent to which the notebook figures as a constituent element in Otto's mental machinery; it is more that in the absence of certain assumptions about how Otto will relate to his notebook it becomes increasingly difficult to treat the notebook as a source of information about the sorts of beliefs that Otto has.

All this, we suggest, alters the way we think about transparency in an active externalist context. Rather than regard transparency as a criterion for cognitive extension, or as something that helps us resolve the constitutive relevance of bio-external resources, our own approach ties transparency to the way we ascribe certain sorts of mental/cognitive dispositional properties to particular cognitive agents. This sort of role is not best understood with respect to the notion of a "constituency condition." Instead, we suggest that issues of transparency are more concerned with ownership-related issues, specifically, the problem of cognitive ownership. This is a recognized problem within the active externalist literature,²¹ although it has seldom been the target of specialist philosophical attention. The problem of cognitive ownership concerns the way we

²¹ As noted by Clark (2011, p. 454), much of the work relating to extended cognition and the extended mind is "best seen as an investigation of... [the] conditions which must be met so as to ensure the *proper ownership* of some candidate extended process by a distinct cognitive agent..." (original emphasis). For detailed treatments of the problem of cognitive ownership, see Wheeler (2018), Rupert (2013), and (especially) Rowlands (2010, Chap. 6).

assign ‘ownership’ of extended cognitive/mental phenomena to a particular cognitive agent, as when we say that a given cognitive capacity (or other dispositional property) ‘belongs’ to a particular human individual. According to the dispositional hypothesis, the problem of cognitive ownership is the problem of ensuring that a given cognitive agent (e.g., a human individual) possesses a certain dispositional property (a cognitive capacity, ability, dispositional belief, and so on). It is precisely at this point, we suggest, that issues of phenomenal transparency start to look important, for transparency is a feature of situations in which matters of cognitive ownership are resolved via our ascriptive efforts, as when we say that agent X believes (in a dispositional sense) that MoMA is on 53rd Street or that agent Y has the ability/capacity to solve long multiplication problems.

By drawing attention to the link between transparency and disposition ascriptions, we avoid many of the problems and pitfalls that were discussed in Sect. 3. In respect of constitutive relevance (see Sect. 3.1), for example, we can see that transparency has absolutely no bearing on the way we individuate the components of mechanisms. At the same time, however, transparency is likely to be a feature of those situations in which we credit ourselves and others with the possession of certain dispositional properties—properties that, when manifest, are realized by the operation of extended mechanisms.

Exotic forms of cognitive extension (see Sect. 3.2) present no problem for the dispositional hypothesis, for the dispositional hypothesis is neutral as regards the nature of E (the subject of cognitive extension). E could, for example, be a plant, a spider, or an AI system, thereby allowing for the possibility of non-human forms of cognitive extension. The extent to which spiders possess any sort of phenomenology is, of course, unclear; nevertheless, if we were to encounter a spider that appeared unable to spin a web, then we would no doubt be reluctant to credit the spider with a web-spinning ability. In this case, there would be no reason to talk of extended abilities (cognitive or otherwise), because the spider just wouldn’t have the abilities that, in more propitious circumstances, might have relied on the instantiation of a mechanism that extended beyond the borders of the spider’s organismic boundary (see Smart et al., 2010).

When it comes to the classic case of Otto and the extended mind, we can see how the dispositional hypothesis obviates the potential tension between phenomenal and functional parity (see Sect. 3.3). From the standpoint of the dispositional hypothesis, the overarching concern is to ensure that a certain sort of dispositional property (i.e., a dispositional belief) is ascribed to Otto. Transparency may well be a feature of such situations—situations in which we are content to make this sort of disposition ascription—but it need not be an essential (i.e., necessary) feature of extended mind scenarios.

Finally, the dispositional hypothesis addresses many of the concerns raised by the shifting boundaries thesis (see Sect. 3.4). Note, for example, that dispositional properties tend to be somewhat enduring. An ability to solve long multiplication problems is thus something that can persist despite the presence of hitches and glitches to the runtime manifestation of that ability. If the pen should break while one is attempting to solve a long multiplication problem, then one’s ability to solve such problems is not lost as a result. To be sure, if one should encounter a lot of problems with the

exercise or manifestation of an ability, then one might begin to wonder if it is still appropriate to regard oneself as possessing that ability (others may wonder about this too). This, however, is perfectly consistent with the idea that transparency is a feature of situations in which matters of cognitive ownership are resolved courtesy of our ascriptive efforts.

5 Conclusion

The active externalist literature suggests that there is an important link between phenomenal transparency and cognitive extension. To date, however, the nature of this link remains obscure. In the present paper, we examined two ways of understanding the appeal to transparency. The first relates to the idea that transparency is necessary for cognitive extension. The second relates to the idea that transparency can be used to resolve disputes pertaining to the constitutive (versus merely causal) status of bio-external resources.

The present analysis offers little in the way of support for either of these proposals. We thus reject the idea that transparency ought to be seen as necessary for cognitive extension. We also challenge the idea that transparency can be interpreted along the lines of a constituency condition for bio-external resources. Inasmuch as transparency has any bearing on constitutional matters, it is, we suggest, unlikely to be one that is best understood via the notion of constitutive relevance.

In the wake of this critique, we sought to outline a new account of cognitive extension—one that helps us understand the putative importance of transparency to claims about extended cognition and the extended mind. According to this account—dubbed the dispositional hypothesis—cognitive extension occurs when the mechanisms responsible for the manifest occurrence of a cognitive/mental dispositional property (e.g., a cognitive ability, cognitive capacity, or dispositional belief) include components that lie beyond the borders of the thing (e.g., the human individual) to which the dispositional property is ascribed. Relative to this way of understanding cognitive extension, transparency is a feature of situations that support the ascription of dispositional properties to both ourselves and others. In cases of extended cognition, for example, transparency looks to be important because it is associated with the proficient use of tools as part of the performance of cognitive tasks. But it is not really the transparency that motivates claims of cognitive extension; it is more that proficiency is one of the things that leads us to credit ourselves and others with the possession of dispositional properties, and it is the possession of these dispositional properties that provides the basis for claims of cognitive extension.

All this, we suggest, alters the way we think about transparency in an active externalist context. Rather than being relevant to constitutional problems (i.e., whether or not a given resource forms part of an individual's cognitive machinery), the notion of transparency is best understood with respect to a somewhat different problem, namely, the problem of cognitive ownership. From the standpoint of the dispositional hypothesis, this problem is concerned with the ascription of dispositional properties—the conditions under which certain entities can be said to possess certain dispositional properties. Transparency looks to be important here because it is a phenomenological

feature of situations that inform the nature of our ascriptive efforts. This is especially so when it comes to our understanding of our own abilities and capacities, and thus our subjective sense of who we are, what we know, and what we can do. Again, however, it is a mistake to regard transparency as a criterion for cognitive extension. Transparency is just a feature of situations in which we are inclined to credit ourselves and others with the possession of certain dispositional properties—properties that, when manifest, rely on forces and factors that lie external to the borders of skin and skull.

Acknowledgements The authors would like to express their thanks to two anonymous reviewers for their comments on an earlier draft of the paper.

Funding This work is partially funded by Portuguese national funds through FCT-Fundação para a Ciência e a Tecnologia under the project UIDB/00183/2020.

Code Availability Not applicable.

Declarations

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this article.

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

- Andrada, G. (2020). Transparency and the phenomenology of extended cognition. *Límite: Interdisciplinary Journal of Philosophy & Psychology*, 15(20), 1–17.
- Andrada, G. (2021). Mind the notebook. *Synthese*, 198, 4689–4708. <https://doi.org/10.1007/s11229-019-02365-9>.
- Andrada, G., Clowes, R. W., & Smart, P. R. (2022). Varieties of transparency: Exploring agency within AI systems. *AI & Society*. <https://doi.org/10.1007/s00146-021-01326-6>.
- Boem, F., Ferretti, G., & Caiani, S. Z. (2021). Out of our skull, in our skin: The microbiota–gut–brain axis and the extended cognition thesis. *Biology & Philosophy*, 36(Article 14), 1–32. <https://doi.org/10.1007/s10539-021-09790-6>.
- Carter, A. J., Clark, A., Kallestrup, J., Palermos, O. S., & Pritchard, D. (Eds.). (2018a). *Extended epistemology*. Oxford University Press.
- Carter, J. A., Clark, A., & Palermos, S. O. (2018b). New humans? Ethics, trust and the extended mind. In A. J. Carter, A. Clark, J. Kallestrup, O. S. Palermos, & D. Pritchard (Eds.), *Extended epistemology* (pp. 331–351). Oxford University Press.
- Clark, A. (2003). *Natural-born cyborgs: Minds, Technologies and the future of human intelligence*. Oxford University Press.
- Clark, A. (2005). Intrinsic content, active memory and the extended mind. *Analysis*, 65(285), 1–11. <https://doi.org/10.1111/j.1467-8284.2005.00514.x>.
- Clark, A. (2008a). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford University Press.

- Clark, A. (2008b). Pressing the flesh: A tension in the study of the embodied, embedded mind? *Philosophy and Phenomenological Research*, 76(1), 37–59. <https://doi.org/10.1111/j.1933-1592.2007.00114.x>.
- Clark, A. (2010). Memento's revenge: The extended mind, extended. In R. Menary (Ed.), *The extended mind* (pp. 43–66). MIT Press.
- Clark, A. (2011). Finding the mind. *Philosophical Studies*, 152(3), 447–461. <https://doi.org/10.1007/s11098-010-9598-9>.
- Clark, A. (2015). What 'extended me' knows. *Synthese*, 192(11), 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/10.1111/1467-8284.00096>.
- Clowes, R. W. (2020). The Internet extended person: Exoself or Doppelganger? *Límite: Interdisciplinary Journal of Philosophy & Psychology*, 15(Article 22), 1–23.
- Craver, C. (2007a). Constitutive explanatory relevance. *Journal of Philosophical Research*, 32, 3–20. <https://doi.org/10.5840/jpr20073241>.
- Craver, C. (2007b). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Clarendon Press.
- Craver, C. F., Glennan, S., & Povich, M. (2021). Constitutive relevance & mutual manipulability revisited. *Synthese*, 199(3), 8807–8828. <https://doi.org/10.1007/s11229-021-03183-8>.
- Facchin, M. (2022). Phenomenal transparency, cognitive extension, and predictive processing. *Phenomenology and the Cognitive Sciences*. <https://doi.org/10.1007/s11097-022-09831-9>.
- Facchin, M., Viola, M., & Zanin, E. (2021). Retiring the Cinderella view: The spinal cord as an intrabodily cognitive extension. *Biology & Philosophy*, 36(Article 45), 1–25. <https://doi.org/10.1007/s10539-021-09822-1>.
- Fazekas, P. (2013). The extended mind thesis and mechanistic explanations. In D. Moyal-Sharrock, V. A. Munz, & A. Coliva (Eds.), *Mind, language, and action* (pp. 125–127). Austrian Ludwig Wittgenstein Society.
- Glennan, S. (2017). *The new mechanical philosophy*. Oxford University Press.
- Glennan, S., & Illari, P. M. (Eds.). (2018a). *The Routledge handbook of mechanisms and mechanical philosophy*. Routledge.
- Glennan, S., & Illari, P. M. (2018b). Varieties of mechanisms. In S. Glennan & P. M. Illari (Eds.), *The Routledge handbook of mechanisms and mechanical philosophy* (pp. 91–103). Routledge.
- Heersmink, R. (2015). Dimensions of integration in embedded and extended cognitive systems. *Phenomenology and the Cognitive Sciences*, 14(3), 577–598. <https://doi.org/10.1007/s11097-014-9355-1>.
- Heidegger, M. (1927). *Being and time*. Basil Blackwell.
- Hurley, S. (2010). The varieties of externalism. In R. Menary (Ed.), *The extended mind* (pp. 101–153). MIT Press.
- Hüttemann, A., & Kaiser, M. I. (2018). Potentiality in biology. In M. Quante (Ed.), *Handbook of potentiality* (pp. 401–428). Springer.
- Illari, P. M., & Williamson, J. (2012). What is a mechanism? Thinking about mechanisms across the sciences. *European Journal for Philosophy of Science*, 2(1), 119–135. <https://doi.org/10.1007/s13194-011-0038-2>.
- Japyassú, H. F., & Laland, K. N. (2017). Extended spider cognition. *Animal Cognition*, 20(3), 375–395. <https://doi.org/10.1007/s10071-017-1069-7>.
- Kaiser, M. I., & Krickel, B. (2017). The metaphysics of constitutive mechanistic phenomena. *The British Journal for the Philosophy of Science*, 68(3), 745–779. <https://doi.org/10.1093/bjps/axv058>.
- Kaplan, D. M. (2012). How to demarcate the boundaries of cognition. *Biology & Philosophy*, 27(4), 545–570. <https://doi.org/10.1007/s10539-012-9308-4>.
- Kirk, R. (2015). Zombies. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2015). Stanford University.
- Kirsh, D. (2019). When is a mind extended? In M. Colombo, E. Irvine, & M. Stapleton (Eds.), *Andy Clark and his critics* (pp. 128–142). Oxford University Press.
- Merleau-Ponty, M. (1945). *Phenomenology of perception*. Routledge Press.
- Miłkowski, M., Clowes, R., Rucińska, Z., Przegalińska, A., Zawidzki, T., Krueger, J., et al. (2018). From wide cognition to mechanisms. *Frontiers in Psychology*, 9(Article 2393), 1–17. <https://doi.org/10.3389/fpsyg.2018.02393>.

- Mumford, S. (2009). Causal powers and capacities. In H. Beebe, C. Hitchcock, & P. Menzies (Eds.), *The Oxford handbook of causation* (pp. 265–278). Oxford University Press.
- Parise, A. G., Gagliano, M., & Souza, G. M. (2020). Extended cognition in plants: Is it possible? *Plant Signaling & Behavior*, 15(2), 1710661. <https://doi.org/10.1080/15592324.2019.1710661>.
- Piredda, G., & Di Francesco, M. (2020). Overcoming the past-endorsement criterion: Toward a transparency-based mark of the mental. *Frontiers in Psychology*, 11, 1–9. <https://doi.org/10.3389/fpsyg.2020.01278>.
- Pritchard, D. (2010). Cognitive ability and the extended cognition thesis. *Synthese*, 175(1), 133–151. <https://doi.org/10.1007/s11229-010-9738-y>.
- Rowlands, M. (2010). *The new science of the mind: From extended mind to embodied phenomenology*. MIT Press.
- Rupert, R. D. (2004). Challenges to the hypothesis of extended cognition. *Journal of Philosophy*, 101(8), 389–428. <https://doi.org/10.5840/jphil2004101826>.
- Rupert, R. D. (2013). Memory, natural kinds, and cognitive extension; or, Martians don't remember, and cognitive science is not about cognition. *Review of Philosophy and Psychology*, 4(1), 25–47. <https://doi.org/10.1007/s13164-012-0129-9>.
- Sims, M., & Kiverstein, J. (2022). Externalized memory in slime mould and the extended (non-neuronal) mind. *Cognitive Systems Research*, 73, 26–35. <https://doi.org/10.1016/j.cogsys.2021.12.001>.
- Smart, P. R. (2018). Human-extended machine cognition. *Cognitive Systems Research*, 49, 9–23. <https://doi.org/10.1016/j.cogsys.2017.11.001>.
- Smart, P. R. (2022). Toward a mechanistic account of extended cognition. *Philosophical Psychology*. <https://doi.org/10.1080/09515089.2021.2023123>.
- Smart, P. R., Engelbrecht, P. C., Braines, D., Strub, M., & Giammanco, C. (2010). The network-extended mind. In D. Verma (Ed.), *Network science for military coalition operations: Information extraction and interaction* (pp. 191–236). IGI Global.
- Thompson, E., & Stapleton, M. (2009). Making sense of sense-making: Reflections on enactive and extended mind theories. *Topoi*, 28(1), 23–30. <https://doi.org/10.1007/s11245-008-9043-2>.
- Wheeler, M. (2010). In defense of extended functionalism. In R. Menary (Ed.), *The extended mind* (pp. 245–270). MIT Press.
- Wheeler, M. (2018). Knowledge, credit, and the extended mind, or what Calvisius Sabinus got right. In A. J. Carter, A. Clark, J. Kallestrup, O. S. Palermos, & D. Pritchard (Eds.), *Extended epistemology* (pp. 127–144). Oxford University Press.
- Wheeler, M. (2019). The reappearing tool: Transparency, smart technology, and the extended mind. *AI & Society*, 34(4), 857–866. <https://doi.org/10.1007/s00146-018-0824-x>.
- Wikforss, Å. (2014). Extended belief and extended knowledge. *Philosophical Issues*, 24(1), 460–481. <https://doi.org/10.1111/phils.12043>.
- Wilson, R. A., & Clark, A. (2009). Situated cognition: Letting nature take its course. In P. Robbins & M. Aydede (Eds.), *The Cambridge handbook of situated cognition* (pp. 55–77). Cambridge University Press.
- Ylikoski, P. (2018). Social mechanisms. In S. Glennan & P. M. Illari (Eds.), *The Routledge handbook of mechanisms and mechanical philosophy* (pp. 401–412). Routledge.
- Zednik, C. (2011). The nature of dynamical explanation. *Philosophy of Science*, 78(2), 238–263. <https://doi.org/10.1086/659221>.