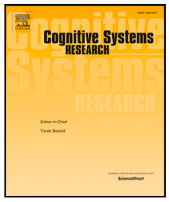
Cognitive Systems Research 84 (2024) 101202

Contents lists available at ScienceDirect

Cognitive Systems Research

journal homepage: www.elsevier.com/locate/cogsys



Extended X: Extending the reach of active externalism

Paul R. Smart

Electronics & Computer Science, University of Southampton, University Road, Southampton, SO17 1BJ, Hampshire, United Kingdom

ARTICLE INFO

Action editor: Serge Thill

Keywords:
Active externalism
Extended cognition
Extended mind
Mechanism
Disposition

ABSTRACT

The terms "extended cognition" and the "extended mind" identify two strands of philosophical argument that are commonly subsumed under the general heading of active externalism. The present paper describes an integrated approach to understanding extended cognition and the extended mind—one that papers over the differences between these two, ostensibly distinct, forms of cognitive extension. As an added bonus, the paper describes how active externalism might be applied to the realm of non-cognitive phenomena, thereby yielding an expansion in the theoretical and empirical scope of the active externalist enterprise. Both these points of progress stem from what is called the *dispositional hypothesis*. According to the dispositional hypothesis, extended cognition occurs when the mechanisms responsible for the manifestation of dispositional properties include components that lie beyond the borders of the thing to which the dispositional properties are ascribed.

1. Introduction

According to the philosophical position known as active externalism, cognitive and mental phenomena can sometimes qualify as extended phenomena in the sense that extra-organismic resources can form part of the causally-active physical fabric that realizes cognitive/mental states and processes (Clark, 2008; Clark & Chalmers, 1998; Menary, 2010). This sort of idea is sometimes presented under the banner of "extended cognition", while at other times, it is presented under the banner of the "extended mind". In both cases, however, the general idea is that extra-organismic resources can, on occasion, become incorporated into the machinery of the mind, such that they are just as much the realizers of mental states and processes as is a neural circuit or a biological brain region. This, at least, is how active externalist theses are presented in the philosophical literature. Consider, for example, the following characterizations of the active externalist position:

> [...] extended mind theory suggests that 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. 3758)

> [...] the actual local operations that realize certain forms of human cognizing include inextricable tangles of feedback, feedforward, and feed-around loops: loops that promiscuously criss-cross the boundaries of brain, body, and world. The local mechanisms of mind, if this is correct, are not all in the head. Cognition leaks out into body and world. (Clark, 2008, p. xxviii)

At first sight, these characterizations can seem unproblematic. This is not to say that there is nothing contentious about the sort of claim that is being made here; it is merely to suggest that the foregoing characterizations yield a seemingly straightforward understanding of what active externalism is all about.

Or do they? It turns out that neither of these characterizations are entirely devoid of problems. The appeal to tangled loops, for example, seems perfectly appropriate for some cases of extended cognizing, but theorists have struggled to reconcile this with the seemingly simple 'loop' that serves as the basis for claims about the extended mind (see Palermos, 2014). Another problem centers on the appeal to "physical machinery" and "local mechanisms". While these mechanistic concepts can be applied to situations involving an occurrent cognitive process; they are much harder to apply to situations involving dispositional kinds, such as states of dispositional belief. Kaiser and Krickel (2017), for example, suggest that mechanistic explanations describe the mechanisms responsible for occurrent phenomena, but occurrent phenomena are not dispositions, and this raises doubts about the extent to which mechanistic concepts can be used to fix the *extended* status of dispositional beliefs.

These points of confusion and ambiguity suggest that we need a clearer explication of active externalism—one that gives us a more precise understanding of what is entailed by the notions of extended cognition and the extended mind. A similar point is made by Chalmers (2019, p. 12) when he suggests that we need a stronger formulation

E-mail address: ps02v@ecs.soton.ac.uk.

https://doi.org/10.1016/j.cogsys.2023.101202
Received 27 March 2022; Accepted 14 December 2023
Available online 2 January 2024

of the extended mind thesis, one that "captures what is really at issue in the debate". Chalmers' own response to this challenge comes in the form of what I will call the *sensorimotor hypothesis*. According to this hypothesis:

> A subject's cognitive processes and mental states can be partly constituted by entities that are external to the subject, in virtue of the subject's sensorimotor interaction with these entities. (Chalmers, 2019, p. 15)

While this formulation helps to address some of the problems that have surfaced in the active externalist literature; it still leaves many questions unanswered. (Consider, for example, that it does not really address the issues raised above.) What is perhaps worse is that the appeal to sensorimotor interaction threatens to introduce additional problems. What is it, for example, that makes some episode of interaction a specifically *sensorimotor* interaction? How do we individuate a sensorimotor interface? And what (if anything) does sensorimotor interaction have to do with the constitutional status of "entities that are external to the subject"?

My aim in the present paper is to present an alternative approach to understanding extended cognition and the extended mind, one that avoids the potentially problematic appeal to sensorimotor interaction. From a methodological standpoint, this account is informed by a consideration of (what I take to be) some puzzling features of the active externalist literature. For the sake of convenience, I will refer to these features under the headings of the duality puzzle, the puzzle of exotic kinds, the puzzle of extended mechanisms, and the missing link puzzle.

The duality puzzle concerns the nature of the relationship between extended cognition and the extended mind. Is the term "extended cognition" merely a terminological variant of the term "extended mind", or do these terms denote two distinct forms of cognitive extension ("cognitive extension" being the term I use to refer to both extended cognition and the extended mind)? In response to this question, the existing philosophical literature reveals a number of ways of distinguishing extended cognition from the extended mind (Allen-Hermanson, 2013; McKenna, 2019; Palermos, 2014; Pöyhönen, 2014; Wheeler, 2019a). Some theorists, for example, have drawn attention to a state vs. process distinction: arguments for the extended mind tend to direct their attention to the realm of cognitive/mental *states*, whereas arguments for extended cognition tend to direct their attention to the realm of cognitive/mental *processes* (e.g., Pöyhönen, 2014). Another distinction relates to the notion of explanatory kinds, with arguments for the extended mind featuring an appeal to explanatory kinds relevant to folk psychology (e.g., belief), and arguments for extended cognition featuring an appeal to explanatory kinds relevant to cognitive science (e.g., memory) (e.g., Allen-Hermanson, 2013). Other sorts of distinction could undoubtedly be made (e.g., arguments for extended cognition tend to focus on *occurrent* cognitive phenomena, while arguments for the extended mind tend to focus on states of *dispositional* belief), but regardless of the way in which the distinction is made, there does appear to be some sort of philosophically-significant difference between the terms "extended cognition" and the "extended mind". Perhaps, then, active externalism is something of a dualistic enterprise—a philosophical fabric spun from two distinct, albeit overlapping, strands of philosophical argument. On the other hand, it would be nice if these two forms of cognitive extension could be subsumed within a common theoretical framework, one that respects the differences between extended cognition and the extended mind, while simultaneously revealing them to be subtly different manifestations of what amounts to the same basic idea. Of these two options, it is the latter, integrative, option that is favored by the theoretical account to be described below.

A second puzzle for the proponent of active externalism relates to the recent emergence of 'exotic' forms of cognitive extension. These include the likes of extended spider cognition (Japyassú & Laland, 2017), extended plant cognition (Parise et al., 2020), extended protist cognition (Sims & Kiverstein, 2022), and the various forms of extended cognizing implemented by non-biological systems, such as Artificial Intelligence (AI) systems (Jonker, 2008; Smart, 2018). For the most part, the bulk of the philosophical literature has focused on human-based forms of cognitive extension—the forms of cognitive extension that involve an appeal to *human* mental/cognitive states/processes. The presence of exotic varieties, however, suggests that the notion of extended cognition (and perhaps the extended mind) might be applicable to non-human entities. At present, it is unclear how these exotic varieties might be accommodated within an overarching theoretical framework that preserves the ideas and insights yielded by a selective focus on human-based forms of cognitive extension. This, then, is a further puzzle for the proponent of active externalism. It is what I will call the *puzzle of exotic kinds*.

A third puzzle is what I will call the *puzzle of extended mechanisms*. In this case, the puzzle relates to the precise role played by mechanistic concepts in active externalist theorizing. Consider, for example, that the active externalist literature features a persistent appeal to mechanism-related concepts, especially talk of extended mechanisms (Clark, 2011; Hurley, 2010; Kaplan, 2012; Smart, 2022; Zednik, 2011). Quite plausibly, the term "extended mechanism" is being used to refer to the mechanisms that are responsible for extended cognitive processes, as well as other extended cognitive phenomena (e.g., extended cognitive states). At present, however, it remains unclear how we ought to understand the notion of an extended mechanism. What is it, exactly, that makes a mechanism extended? Presumably, the answer has something to do with the fact that the mechanism transcends some sort of border or boundary, but the nature of that border or boundary is somewhat hard to pin down. Active externalists typically appeal to biological or metabolic boundaries, such as the proverbial borders of "skin and skull", in referring to extended mechanisms. But this sort of characterization will not work for the more exotic forms of cognitive extension mentioned above. Plants, for example, do not have brains, spiders do not have skulls, and AI systems do not have skin. Part of the problem here is that different kinds of cognitive extension feature entities/agents with different borders and boundaries, so our understanding of what makes something an extended mechanism must be pitched at a level of abstraction that is divorced from the material features of any particular entity or agent.

A final puzzle relates to the historical precursors of active externalism. Active externalism is mostly directed to the realm of cognitive/mental phenomena (e.g., cognitive processes, dispositional beliefs, and so on). Historically, however, active externalist claims emerged against a backdrop of cases that did not involve an appeal to cognitive/mental phenomena. The swimming-related capabilities of bluefin tuna are a nice example of this (see Clark, 1997; Kaplan, 2012). It is hard to see how such capabilities could be characterized in cognitive terms, and yet the explanation of the tuna's natatorial success is one that bears a striking resemblance to the explanatory approach adopted for cases of extended cognizing (see Section 4.3). What is more, the proponents of active externalism sometimes appeal to non-cognitive cases as a means of explicating active externalist ideas. Wilson and Clark (2009, p. 62), for example, refer to a putative form of (non-cognitive) 'extension' that centers on hermit crabs. They also draw attention to the parallels between extended cognition and extension-related concepts in disciplines as diverse as computer science (extended computation), evolutionary biology (niche construction), physiology (extended physiology), and developmental systems theory. All this presents us with another puzzle. Call it the *missing link puzzle*. This puzzle concerns the relationship between active externalism – as a philosophical movement within the sciences of the mind – and the forms of extension that occur across a broad swath of other disciplines, some of which lie beyond the borders of cognitive science (see Huneman, 2013).

The present paper aims to resolve all these puzzles by describing a theoretical account dubbed the dispositional hypothesis (see Section 3). This hypothesis represents a generalization of another hypothesis that

The Cognitive Capacity Hypothesis

[CC1] A cognitive capacity (CC) is ascribed to a given human individual (H).

[CC2] The exercise of CC is a cognitive process (CP).

[CC3] CP is realized/constituted by a mechanism (M).

[CC4] The constituents of M are a set of components (A).

[CC5] In the case of *non-extended cognition*, the members of A are wholly contained within the borders/boundaries of H.

[CC6] In the case of *extended cognition*, some of the members of A are located external to the borders/boundaries of H.

Fig. 1. The Cognitive Capacity Hypothesis.

was developed to cater for cases involving an appeal to extended cognitive processes (see Section 2). Courtesy of the generalization, I show how the dispositional hypothesis provides us with a relatively straightforward approach to understanding extended cognition and the extended mind (see Section 4.1). I also show how the dispositional hypothesis applies to both cognitive and non-cognitive phenomena (see Section 4.3), thereby expanding the scope of active externalist theorizing to the realm of Extended X,[1] (where the "X" refers to phenomena of both the cognitive and non-cognitive kind).

2. The cognitive capacity hypothesis

In confronting a complex problem, it sometimes helps to limit one's attention to a single, simplified version of the problem and then generalize the solution (if there is one) to a wider class of cases. This is the strategy I will employ here. I will thus begin by considering the case of long multiplication, which is a frequently cited example of extended cognizing within the active externalist literature (Wheeler, 2010; Wilson & Clark, 2009). The case centers on the use of pen and paper resources to solve long multiplication problems, such as 763×342. I will not detain the reader by discussing the details of this case, in part because many readers will be familiar with the relevant problem-solving routine. In a nutshell, when we are presented with a long multiplication problem, we often resort to a problem-solving strategy that involves the use of pen and paper resources. These extra-organismic resources are used as part of the multiplicative process, and, if everything goes according to plan, the process terminates in the correct solution to the original problem. This is what we might call a form of extended problem-solving (Kirsh, 2009) or (perhaps better) a form of extended mathematical cognition (Menary, 2015).

For the purposes of the present analysis, let us assume that the long multiplication case – the use of pen and paper resources to solve long multiplication problems – counts as a *bona fide* example of extended cognizing. The analysis of this case yields a theoretical account of extended cognition that I will dub the *cognitive capacity hypothesis* (see Fig. 1).

The cognitive capacity hypothesis assumes that a cognitive capacity (CC) is ascribed to a particular individual (clause CC1).[2] In the long multiplication case, the relevant capacity is a capacity to solve long multiplication problems. This is what we might call a mathematical capacity or, more specifically, a multiplicative capacity. For other forms of extended cognizing, the ascribed capacity will be of a different kind. In the case of extended memory (Carter & Kallestrup, 2016), for example, the ascribed capacity is likely to be a mnemonic capacity, while in the case of extended perception (Wilson, 2010), the ascribed capacity is likely to be a perceptual capacity.[3]

The exercise of CC is glossed as a cognitive process (CP) (clause CC2). For the long multiplication case, CP corresponds to the actual long multiplication process, i.e., the occurrent process that involves the use of pen and paper resources to solve the long multiplication problem. The claim that this particular process reflects the exercise of some sort of capacity is, I think, largely uncontroversial. What is perhaps more controversial is the idea that this process ought to be regarded as a specifically cognitive process. Does the world-involving variant of the long multiplication process count as a genuinely cognitive process, and, if so, what is it that makes it a cognitive process?

In response to this question, it is worth noting that if the long multiplication process were to be performed in the head, using nothing more than the computational and representational resources of the biological brain, then we would probably have little problem in accepting the cognitive status of the long multiplication process.[4] In this

[1] The term "Extended X" is owed to Huneman (2013).

[2] There ought to be nothing controversial about this appeal to capacities as the starting point for our understanding of extended cognition. As noted by Cummins (2000), capacities play an important role in scientific efforts to explain a broad array of psychological phenomena: "The primary explananda of psychology [are] capacities: the capacity to see depth, to learn and speak a language, to plan, to predict the future, to empathize, to fathom the mental states of others, to deceive oneself, to be self-aware, and so on. Understanding these sorts of capacities is what motivates psychological inquiry in the first place" (Cummins, 2000, p. 122).

[3] This highlights the generality of the cognitive capacity hypothesis relative to the various forms of extended cognition that have been discussed in the active externalist literature. Thus, while the cognitive capacity hypothesis is inspired by a consideration of one particular form of extended cognition, namely the use of pen and paper resources to solve long multiplication problems, it should not be seen as limited to the long multiplication case. The cognitive capacity hypothesis is intended to serve as the basis for a general account that applies to multiple forms of extended cognizing, but the ultimate endpoint of this analytic trajectory is the dispositional hypothesis, which is discussed in Section 3.

[4] This represents an application of what has been called the parity principle (see Clark & Chalmers, 1998).

respect, the long multiplication case is interesting, for there are indeed times when we resort to the use of in-the-head methods to solve long multiplication problems. Consider, for example, that if we were pressed to solve the problem 763 × 342 without the use of pen and paper resources, then we could probably do so via an in-the-head routine. The upshot is that the ascription of a multiplicative capacity comes with a degree of uncertainty. In ascribing a multiplicative capacity to a human individual, we are assuming that the individual has a capacity to solve long multiplication problems. But the exercise of this capacity is indeterminate between the use of in-the-head and in-the-world methods. We may believe that a given individual will solve long multiplication problems in-the-head, but then discover, to our surprise, that these problems are being solved in-the-world. Conversely, we may believe that an individual will solve long multiplication problems using pen and paper resources, but then discover that they resort to an in-the-head strategy. Inasmuch as we are content to assume that the ascribed capacity counts as cognitive when the exercise of this capacity involves the in-the-head strategy, then what reason do we have to discount the cognitive status of (what looks to be) the same capacity when the exercise of the capacity relies on the deployment on some alternative problem-solving strategy (i.e., an in-the-world strategy)? After all, prior to the observation of token instantiations of the multiplicative process, we may not know how the capacity will be exercised. And even if we did, would this materially alter our intuitions about the cognitive status of the sort of capacity that is being ascribed here?

This highlights an important point about the cognitive capacity hypothesis: The hypothesis assumes that the cognitive status of a process (the exercise of a capacity) stems from the cognitive status of whatever capacity the process is the exercise of. Accordingly, in the long multiplication case, the idea is that the long multiplication process counts as cognitive simply because it corresponds to the exercise of a capacity that we are content to regard as cognitive. This capacity-centric approach to the problem of cognitive status is to be contrasted with an alternative approach that directs attention to the features of cognitive processes and/or the mechanisms responsible for these processes. This alternative approach is well-represented in theoretical debates pertaining to the so-called "mark of the cognitive" (Adams, 2010; Adams & Garrison, 2013).[5]

As will become clear in the next section, I regard capacities as being a subset of the class of dispositions. A capacity describes the powers of some object (in this case, a human individual) to bring about a certain state-of-affairs (e.g., the solution to long multiplication problems). While the notion of a disposition might be seen to refer to things that will happen in the future, there is, I think, no reason why we could not (and, indeed, would not) ascribe a disposition based on the observation of a specific cognitive performance. If, for example, we observe a human individual expertly solving long multiplication problems using pen and paper resources, then it is reasonable to assume that the individual has a capacity to solve long multiplication problems. And this is so, even if prior to the observation of such performances we were unsure as to whether the individual possessed this capacity. Such ascriptions of skill and expertise no doubt play an important role in enabling us to rely on individuals in particular circumstances. At the very least, they enable us to anticipate the behavior of individuals in future to-be-encountered situations. Thus, if we were to ever rely on an individual to solve long multiplication problems, we would no doubt want to be sure that the individual could, in fact, solve such problems. We would, in short, want to know that we could rely on the individual to do the things we expected them to do.

In the world-involving variant of the long multiplication process, we ascribe a capacity to a human individual, but the exercise of this capacity (the long multiplication process) is one that features the use of resources that lie external to the human individual (i.e., the process involves resources that lie external to the borders of the thing to which the capacity is ascribed). Does this mean our capacity-related ascriptions are misplaced? Should we perhaps say that it is the larger system, consisting of human + pen + paper, that has the multiplicative capacity rather than the human individual? The answer to this question is, I think, unclear. Nevertheless, for most practical purposes, there are a number of reasons why we might be inclined to regard the human individual as the subject, bearer, or owner of the multiplicative capacity. For a start, it is the human individual that brings the long multiplication routine into existence by creating (or at least triggering the creation of) the mechanism that is responsible for that routine. In addition, the individual is exerting considerable control over the way in which the long multiplication process unfolds. In particular, any individual who possesses a genuine multiplicative capacity will adhere to a set of normative constraints and practices that govern the overall shape of the long multiplication routine (Menary, 2015; Roberts, 2012). (Such constraints are, of course, required to ensure that the long multiplication routine terminates in a successful solution.) In this sense, it seems perfectly appropriate to credit the individual with a capacity to solve long multiplication problems. If, for example, the process succeeds, then we are likely to see the human individual as being 'responsible' for this outcome (see Roberts, 2012). Conversely, if the process should go awry, then it hardly seems appropriate to blame the pen for the anomalous result. Whatever the outcome of the long multiplication process, it is likely to be the human individual that is credited (or blamed) for the success (or failure) of the long multiplication process. (This is what I mean by the individual being 'responsible' for the outcome of the long multiplication process.)

Clauses CC3 and CC4 of the cognitive capacity hypothesis feature an appeal to mechanistic concepts, specifically the concepts of *mechanism* and *component*. These concepts are taken from the burgeoning literature on the philosophy of mechanisms—a field of philosophical research known as (the new) mechanical (or neo-mechanical) philosophy (Glennan, 2017; Glennan & Illari, 2018a). A concise characterization of the mechanism concept is provided by Glennan (2017):

> A mechanism for a phenomenon consists of entities (or parts) whose activities and interactions are organized so as to be responsible for the phenomenon. (Glennan, 2017, p. 17)

As should be clear from this characterization, the constituents of mechanisms are entities and activities. These are what are known as the *components* of mechanisms. Specifically, a component is a composite of both an entity *and* an activity. There are, as Glennan (2017, pp. 20–21) notes, no entities without activities (entities must have activities in order to qualify as components), nor are there any activities without entities (activities must belong to some entity). In this sense, a component is some entity involved in an activity. In the long multiplication case, the set of components includes (I suggest) the human individual, as well as the bio-external pen and paper resources. These components interact in such a way as to be collectively responsible for some phenomenon, which, in the long multiplication case, is the long multiplication process. Together, then, the components comprise the mechanism (they are the mechanism's parts), and the mechanism is responsible for the long multiplication process.

[5] In one sense, of course, this approach to resolving the cognitive status of CP does not really get us anywhere, for we still need some means of determining what it is that makes a given capacity a specifically cognitive capacity. For present purposes, I will assume that we have no problem in identifying certain capacities as cognitive capacities, regardless of our understanding of the precise details of the processes (and, crucially, the underlying mechanisms) that reflect the exercise of these capacities. This makes sense, since I assume that for much of human history, we had little insight into the nature of the mechanisms that realized cognitive processes. And yet, despite this, we were, I presume, still able to recognize certain capacities as being of the cognitive kind.

The sense in which a mechanism is seen to be responsible for a phenomenon can be interpreted in one of two ways. A mechanism can either be seen to cause a phenomenon or it can be seen to constitute a phenomenon (see, for example, Kaiser & Krickel, 2017). As suggested by CC3, the relevant form of responsibility in the case of the cognitive capacity hypothesis is the constitutive one. In particular, a mechanism (M) is deemed to realize/constitute a cognitive phenomenon (CP), or, inversely, CP is deemed to be realized by/constituted by M. The relevant form of realization here is what is called *mechanistic realization*, which, according to Wilson and Craver (2007), is a particular kind of explanatory realization relation. This form of mechanistic realization is, I suggest, semantically equivalent to the notion of *mechanistic constitution* (see Baumgartner et al., 2020), which is the form of constitution also referred to in CC3.[6]

The cognitive capacity hypothesis is broadly consistent with the way that cognitive extension has been discussed in the philosophical literature. Here, for example, is how Clark (2015) refers to the extended mind:

> The point of the extended mind story was to show that, considered in the context of an active, cognitively well-endowed organism, certain apparently bodily or worldly goings-on might form parts of the realization base for some cognitive capacities. (Clark, 2015, p. 3771)

Clark is evidently talking about the extended mind here, which may not be the same as extended cognition. Nevertheless, this quotation does appear to express a view that is nicely aligned with the cognitive capacity hypothesis. Firstly, we have the appeal to cognitive capacities. Secondly, there is the appeal to "worldly goings-on", which I interpret as the activity of entities that lie external to the individual. Finally, there is the appeal to realization. Clark does not express a commitment to a particular form of realization here; nevertheless, I suspect that the notion of mechanistic realization is at least a plausible candidate for the kind of realization that is being suggested.

A potential point of disagreement between Clark and the cognitive capacity hypothesis is likely to arise in respect of the thing to which capacities are ascribed. The cognitive capacity hypothesis suggests that these capacities are ascribed to a human individual, but I suspect Clark would be inclined to regard these capacities as the properties of the larger materially-hybrid cognitive organization (the extended cognitive system) that includes both the human (biological) individual and the resources that lie external to the human individual. Consider, for example, the way that Wilson and Clark (2009) refer to the role of so-called "coupling conditions" in arguments for the extended mind:

> 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). (Wilson & Clark, 2009, p. 67)

My own view, here, is that the capacities belong to the human individual, as opposed to the larger, hybrid system. That, however, is by-the-by. What matters, for present purposes, is not that capacities are being ascribed to the *right* target; what matters is that capacities are, as a matter of fact, being ascribed to one of the constituent elements of a larger systemic organization. What matters, in other words, is that we attribute capacities (rightly or wrongly) to the human individual rather than to the extended cognitive system that actually performs the long multiplication process. This is, I think, compatible with the idea that capacities "are plausibly seen as the capacities of a specific individual" in the above quotation. Of course, given that the larger extended cognitive system in the long multiplication case is a *transient* construction – what Wilson and Clark (2009) dub a Transient Extended Cognitive System (TECS) – we may have little choice but to ascribe the relevant (multiplicative) capacity to the human individual. The reason for this stems from the very notion of a TECS. Given that the extended cognitive system in the long multiplication case is, by its very nature, a *transient* construction, it will only exist when the relevant capacity is exercised (i.e., a token instantiation of the long multiplication routine is performed). For much of the time, then, the extended cognitive system will not exist as a discernible physical system to which any sort of capacity could be 'attached'. The upshot is that there is probably little to be gained (in a pragmatic sense) by ascribing capacities to the larger extended system, consisting of the human individual + bio-external props, aids, and artifacts. Instead, when it comes to the ascription of capacities (to both ourselves and others), we simply see the human individual as possessing these capacities. That is to say, we see the human individual as the subject, bearer, or owner of capacities, even if the exercise of those capacities involves the temporary construction of a larger system that performs a given cognitive process. This arguably makes a great deal of sense, for (as noted above) the mechanism that realizes the exercise of a cognitive capacity is typically one that is brought into existence by the human individual (the entity to which the capacity is ascribed). Thus, in the long multiplication case, it is the human individual that instantiates the mechanism that then realizes the long multiplication process, and it is only when the human individual is willing to engage in the relevant routine that the routine stands any chance of being completed (or, of course, started).

3. The dispositional hypothesis

The cognitive capacity hypothesis caters for the features of at least one instance of extended cognizing that has been discussed in the active externalist literature (i.e., the long multiplication case). As things stand, however, it does not provide us with a means of accommodating non-cognitive phenomena; nor does it tell us much about the relationship between extended cognition and the extended mind. Given that at least one of these problems (the former) relates to the *specificity* of the cognitive capacity hypothesis, it may help to generalize some of the terms used in the cognitive capacity hypothesis. The terms I focus on here are CC, H, and CP.

The first generalization concerns cognitive capacities (denoted by the term CC). I deem cognitive capacities to be members of the class of capacities, which are, in turn, members of the class of things called *dispositions*.[7] Accordingly, in generalizing the cognitive capacity hypothesis, I will substitute the notion of a cognitive capacity with the more generic notion of a disposition or dispositional property.[8]

As noted by Mumford (1998), the term "disposition" subsumes things like abilities, capacities, capabilities, proclivities, powers, potentialities, tendencies, and so on. This raises a worry about the over-generalization of CC1. In particular, it is unclear whether active externalist claims are best understood with respect to the generic class of

[6] For reasons of space, I will not attempt to detail the reasoning behind this inference of semantic equivalence. I will simply assume, without further argument, that these concepts are equivalent.

[7] The claim that capacities are members of the class of dispositions is consistent with the work of a number of theorists (e.g., Cartwright, 2007; Mumford, 1998). According to Glennan (2017, p. 51), "Capacities... are just dispositional properties of systems". Cummins (2000, p. 122) also suggests that capacities are a particular form of dispositional property when he writes that: "Capacities are best understood as a kind of complex dispositional property".

[8] The status of dispositions as dispositional properties is an issue that has been the source of considerable controversy in the philosophical literature (see, for example, Mumford, 1998, chap. 1). For present purposes, I will assume that it is appropriate to talk of dispositional properties. This is consistent with the way that dispositions have been conceptualized in the philosophy of science (e.g., Hüttemann & Kaiser, 2018). For some resistance to the idea that dispositions ought to be regarded as properties, see Mumford (2009).

dispositional properties, or whether such claims ought to be restricted to a particular subset of such properties. While this is an important issue – and one that warrants further attention – I will seek to minimize the number of constraints that are applied to the notion of a dispositional property. The only constraint I will impose relates to that mandated by the appeal to mechanistic concepts in CC3 and CC4. Accordingly, I suggest that the kind of dispositional properties we are interested in are those whose exercise/manifestation involves the instantiation of a mechanism. In other words, in talking about dispositional properties, I will assume that the manifestation of such properties is subject to mechanistic realization/constitution. These are what might be called *mechanism-dependent dispositions*.[9]

The second generalization relates to H (the human individual). In the context of the cognitive capacity hypothesis, this term denotes a human individual. But this emphasis on human individuals merely reflects the peculiar features of the long multiplication case—the fact that it is a human individual that is performing the long multiplication task. Accordingly, let us generalize H to include anything that could be the subject of cognitive extension. These are what I will call *entities*. The choice of terminology here is motivated by the terminological conventions used in neo-mechanical philosophy, which, recall, is the source of other mechanism-related concepts referred to by the cognitive capacity hypothesis.[10] For present purposes, the term "entity" means something like a physical object. It is, in short, an umbrella term for any object to which a (mechanism-dependent) dispositional property might be ascribed (e.g., a system, an agent, a human individual, a biological organism, and so on).[11]

The third, and final, generalization relates to the notion of a cognitive process (denoted by the term CP). I deem cognitive processes to be members of the class of processes, which are, in turn, members of the class of things called *occurrents*. From a metaphysical standpoint, the notion of an occurrent subsumes things like processes, states, and events,[12] all of which are the sorts of things that might be analyzed by scientists (and synthesized by engineers). In one sense, then, the term "occurrent" seems to be a good substitute for the notion of a cognitive process. On the other hand, clause CC3 of the cognitive capacity hypothesis refers to the role of mechanisms in constituting or realizing a cognitive process. This, as noted in Section 2, draws attention to a particular kind of relational construct, namely, the notion of mechanistic realization (Wilson & Craver, 2007) and/or mechanistic constitution (Baumgartner et al., 2020). In short, we want to substitute CP with something that is compatible with the relata of the mechanistic realization/constitution relation, such that it makes sense to say that a mechanism (M) realizes/constitutes Y, where Y refers to whatever it is that is being realized/constituted.

According to Kaiser and Krickel (2017), the nature of what I am calling Y are constitutive mechanistic phenomena or object-involving occurrents, where an object-involving occurrent is glossed as "an object (or system) that is engaged in a certain occurrent" (Kaiser & Krickel, 2017, p. 768). Accordingly, I will generalize the notion of a cognitive process to include anything that could qualify as a constitutive mechanistic phenomenon (or object-involving occurrent).

The generalizations mentioned above lead to the following substitutions of the terms used in the cognitive capacity hypothesis:

- $H \Rightarrow E$ (human individual ⇒ entity).
- $CC \Rightarrow D$ (cognitive capacity ⇒ dispositional property).
- $CP \Rightarrow P$ (cognitive process ⇒ constitutive mechanistic phenomenon or object-involving occurrent).

The upshot is a generalized version of the cognitive capacity hypothesis, which I will call the *dispositional hypothesis* (see Fig. 2).

The dispositional hypothesis suggests that we observe a case of extended cognizing whenever the mechanisms responsible for the manifestation or exercise of a cognitive dispositional property (such as a cognitive capacity) are ones that extend beyond the borders or boundaries of the thing to which the dispositional property is ascribed.[13] These mechanisms are what we might call extended (Clark, 2011; Hurley, 2010; Kaplan, 2012; Smart, 2022; Zednik, 2011), wide (Miłkowski et al., 2018), or supersized (Clark, 2008) mechanisms. (For the sake of simplicity, I will use the term "extended mechanism".) We have thus arrived at a potential resolution of one of the puzzles mentioned in Section 1, namely, the puzzle of extended mechanisms. According to the dispositional hypothesis, a mechanism is judged to be extended whenever it transcends the border or boundary of the entity to which a particular dispositional property is ascribed.

Having now presented the dispositional hypothesis, let us proceed to evaluate the hypothesis. The next section (Section 4) seeks to apply the dispositional hypothesis to a number of cases involving extended (cognitive/mental) phenomena. It also seeks to test whether the dispositional hypothesis is able to distinguish between extended cognition and other, ostensibly similar, forms of cognition, such as distributed cognition (see Section 4.4) and embedded cognition (see Section 4.5).

4. Evaluating the dispositional hypothesis

4.1. Extended minds

One of the objectives of the dispositional hypothesis is to tackle the seemingly dualistic nature of the active externalistic enterprise—the fact that extended cognition and the extended mind are treated as distinct, albeit inter-related, forms of cognitive extension. This objective is achieved courtesy of the appeal to dispositional properties. Such properties subsume the notion of dispositional beliefs, which lie at the heart of the most well-known philosophical exemplar of the extended mind, namely, the Otto notebook case (Clark & Chalmers, 1998). In describing the Otto notebook case, Clark and Chalmers (1998) suggest that the notebook serves as part of the (mechanical) supervenience base for some of Otto's *dispositional* beliefs, e.g., the belief that The Museum of Modern Art (MoMA) is located on 53rd Street. This is what is leading us to the idea that Otto's dispositional beliefs ought to be regarded as *extended* beliefs. We can understand this appeal to extended dispositional beliefs in precisely the same way as we might understand the appeal to extended cognitive capacities in the long multiplication case. In both cases, we are ascribing a dispositional property to a

[9] It is perfectly possible that additional constraints will need to be imposed on dispositional properties, or at least the situations in which it is appropriate to talk of dispositional properties being subject to extended mechanistic realization. Two such constraints are what I will dub the *causal constraint* (the target of disposition ascription – the entity to which dispositional properties are ascribed – should, via their own behavior, play a causal role in instantiating M) and the *inclusivity constraint* (the target of disposition ascription should qualify as a component in M). For reasons of space, I will refrain from further discussion of these constraints.

[10] In fact, the terminological conventions within mechanical philosophy are somewhat vexed. In respect of the term "entity", for example, Glennan (2017, p. 20) writes: "The term 'entity' is the vaguest, and in this sense may be the best, but it has one decided disadvantage: within metaphysics the term 'entity' is used to refer generically to any member of the ontological zoo—so among the entities we might believe in are events, substances, properties, processes, tropes, and so on. It is thus a far broader category than the New Mechanist's entities. But since the use of 'entity' has now become deeply entrenched in the mechanisms literature, I will accede to that usage".

[11] The term "disposition carrier" may be a less metaphysically-loaded way of referring to the subject of extension.

[12] According to Kaiser and Krickel (2017, p. 768), "occurrents are process[es], event[s], and states (where activities, behaviors, and the like are, plausibly, special kinds of processes, events, or states)".

[13] These are what Mumford (1998, p. 1–2) calls *disposition ascriptions*. They are the "attributions of dispositions to individual objects".

The Dispositional Hypothesis

[**DH1**] A dispositional property (D) is ascribed to an entity (E).

[**DH2**] The exercise/manifestation of D is a constitutive mechanistic phenomenon (P).

[**DH3**] P is realized/constituted by a mechanism (M).

[**DH4**] The constituents of M are a set of components (A).

[**DH5**] In the case of extended phenomena (P_X), some of the members of A are located external to the borders/boundaries of E.

[**DH6**] In the case of extended cognitive phenomena (P_{XC}), D qualifies as a cognitive/mental property (e.g., a cognitive capacity or a dispositional belief).

Fig. 2. The Dispositional Hypothesis.

particular individual, but the runtime mechanisms that realize the manifestation of this property are ones that include components that lie external to the borders/boundaries of this individual.

The upshot is that we have effectively resolved the duality puzzle: the dispositional hypothesis is just as applicable to cases featuring an appeal to the extended mind as it is to cases featuring an appeal to extended cognition. The main difference here relates to the nature of the dispositional property that is being ascribed to a particular cognitive individual. In the case of the extended mind, the dispositional properties are drawn from the vocabulary employed by folk psychology (e.g., dispositional beliefs), whereas in the case of extended cognition, the dispositional properties are drawn from the vocabulary employed by cognitive science (e.g., cognitive capacities/abilities). This difference is not necessarily unimportant or insignificant, but there is no reason why a simple shift in the nature of a dispositional property would materially alter the way we understand extended cognition and the extended mind from the standpoint of the dispositional hypothesis.

4.2. Exotic forms of cognitive extension

Thus far, we have seen how the dispositional hypothesis caters for cases involving an appeal to either extended cognition or the extended mind. But the various forms of generalization implemented in Section 3 were only partly geared to accommodating the extended mind. Another objective was to broaden the scope of the active externalist enterprise—to take the enterprise beyond the narrow confines of human-centered cognizing, and, indeed, beyond the confines of cognitive science.

As a means of evaluating the extent to which we have achieved this objective, it will be useful to consider the extent to which the dispositional hypothesis can be applied to *non-human* forms of cognitive extension, i.e., forms of cognitive extension in which the entity E is not a human individual. One example of this stems from recent work in computer science, especially work that seeks to expand the capacities of conventional computational systems by incorporating human individuals into computational routines (e.g., Law & von Ahn, 2011). Such work serves as the basis for what is dubbed human-extended machine cognition (Smart, 2018), an unusual form of extended cognizing in which one or more human individuals are incorporated into the (runtime) cognitive–computational processing loops of a technologically-advanced system, such as an AI system. Human-extended machine cognition is thus a particular form of what might be called *machine-centered extended cognition*, a form of cognitive extension that includes the likes of extended AI (Jonker, 2008).[14]

Such forms of cognitive extension are easily accommodated by the dispositional hypothesis. The reason for this is that the dispositional hypothesis makes no claim about the nature of the entity that is subject to some form of cognitive extension. In philosophical circles, the main target of analytic attention is, of course, human-centered extended cognition—the forms of cognitive extension that are spun around a single human individual. There is, however, nothing about the dispositional hypothesis that excludes the possibility of non-human forms of cognitive extension. The dispositional hypothesis is thus just as applicable to technological systems as it is to human individuals. The same is true of those forms of cognitive extension that are based around a non-human biological entity. Examples include the likes of extended spider cognition (Japyassú & Laland, 2017) extended plant cognition (Parise et al., 2020), and extended cognition in slime molds (Sims & Kiverstein, 2022). In this sense, we have resolved the exotic kinds puzzle: the dispositional hypothesis applies just as well to spiders, plants, and computational systems as it does to the more traditional (human) targets of the active externalist enterprise.

Not everyone, of course, will be happy with the idea of human-extended machine cognition, especially given the computational nature of both the central entity (the AI system) and the information processing routines that reflect the exercise of that entity's (cognitive–computational) capacities. Within philosophical circles, at least, there seems to be considerable resistance to the idea that computational

[14] The notion of extended AI is, of course, not limited to situations where individual humans – the likes of you and me – are incorporated into a cognitive–computational routine. In principle, there is nothing that would prevent the term "extended AI" being applied to situations in which AI systems exploit a surrounding penumbra of *non-human* resources for the completion of cognitive/computational tasks. One example of this stems from recent work into so-called differentiable neural computers. As discussed by Clark (2019, p. 272), these are "deep learning networks that have learnt to use read-write operations to couple their own internal processing capacities to stable yet modifiable external data stores so as to deliver brand new kinds of functionality". Such systems, I suggest, are candidate cases of extended AI, even though human individuals do not serve as components of the relevant extended mechanism (as per the notion of human-extended machine cognition).

systems (especially those trading in the manipulation of symbolic representations) ought to be seen as *bona fide* cognitive entities (see van Gelder, 1995). For present purposes, however, we can park this issue, for nothing about the dispositional hypothesis requires us to make a firm distinction between the realms of the cognitive and the computational (or, indeed, between the realms of the cognitive and the non-cognitive). The dispositional hypothesis is intended to apply to *any* kind of (mechanism-dependent) dispositional property, no matter its cognitive status. Accordingly, we could accept the idea that no computational system – including a human computation system – ought to be regarded as a *bona fide* cognitive system. By itself, however, this will not materially alter claims about the *extended* status of certain forms of computational system (including AI systems).[15] We can thus accept the possibility of extended AI without becoming overly embroiled in the ongoing debate about the seemingly elusive "mark of the cognitive" (see Adams, 2010; Adams & Garrison, 2013).

The dispositional hypothesis can also be put to work in helping us understand intra-bodily forms of cognitive extension, such as those centered on a specific biological organ (e.g., the biological brain) (Boem et al., 2021; Facchin et al., 2021). Understanding these forms of cognitive extension from the perspective of the sensorimotor hypothesis (Chalmers, 2019) is complicated by the fact that the mechanisms of interest do not extend beyond the bio-corporeal boundaries of the human subject. This makes it difficult to identify a specific sensorimotor or perceptuo-motor interface by which the extended status of a cognitive routine could be determined.

In contrast to the sensorimotor hypothesis, intra-bodily forms of cognitive extension present no problem for the dispositional hypothesis. As noted above, the dispositional hypothesis is neutral as regards the nature of the entity to which a dispositional property is ascribed. Accordingly, there is nothing to prevent the dispositional hypothesis being applied to situations where we ascribe a given cognitive capacity to (e.g.) the biological brain, but then discover that the mechanisms responsible for the manifestation of this capacity are ones that extend beyond the neurological realm to include a diverse array of extra-neural (albeit still intra-bodily) resources. In this sense, the dispositional hypothesis is just as applicable to intra-bodily (or sub-personal) forms of extended cognition as it is to those forms of extended cognition that are individuated at the personal or organismic level.

4.3. Extended swimming

In addition to being neutral about the subject of cognitive extension, the dispositional hypothesis is also neutral about the cognitive/non-cognitive status of the phenomena that are deemed to be extended. This neutrality is important, for one of the aims of the hypothesis is to illuminate the nature of the missing link—to extend the remit of active externalist theorizing to the realm of non-cognitive phenomena. We therefore want to establish a sensible point of contact with work in a number of disciplines (most notably the life sciences), all of which have circled around the general idea of extended realization bases for particular kinds of phenomena (see the discussion in Wilson & Clark, 2009).

With this in mind, let us consider a case in which the non-cognitive status of a putatively extended routine ought not to be in any doubt. The case I will focus on here concerns the swimming-related performances of certain marine species, especially the bluefin tuna. The details of this case are described by Clark (1997) and Kaplan (2012). For present purposes, however, I will help myself to the summary provided by Clark (2008):

> The extraordinary efficiency of the fish as a swimming device is partly due, it now seems, to an evolved capacity to couple its swimming behaviors to the pools of external kinetic energy found as swirls, eddies, and vortices in its watery environment (see Triantafyllou & Triantafyllou, 1995). These vortices include both naturally occurring ones (e.g., where water hits a rock) and self-induced ones (created by well-timed tail flaps). The fish swims by building these externally occurring processes into the very heart of its locomotion routines. The fish and surrounding vortices together constitute a unified and remarkably efficient swimming machine. (Clark, 2008, p. 225–226)

In order to apply the dispositional hypothesis to the case of extended swimming, we simply need to ignore the final clause (i.e., DH6) of the dispositional hypothesis (this clause, recall, is only applicable to phenomena of the cognitive kind). Apart from this, however, the dispositional hypothesis is perfectly able to accommodate the bluefin tuna case. In effect, what we are doing is ascribing a certain capacity (a dispositional property) to the bluefin tuna (e.g., a capacity to swim at a certain speed or to swim with a certain efficiency). This capacity is probably not one that we are prepared to accept as cognitive—it is more akin to a physical capacity than it is a cognitive capacity. Despite this, however, there is no reason why we should not seek to provide a mechanistic explanation of this capacity, just as we do with any number of other dispositional properties to be found within the biological sciences (see Hüttemann & Kaiser, 2018). And, once we embark on this mechanistically-oriented explanatory effort, we may discover that our empirical nets need to be cast much more widely than the tuna's organismic boundary. That is to say, as part of our efforts to explain the phenomenon of aquatic locomotion, we may discover that the borders and boundaries of the relevant locomotory mechanism are not quite where we thought they were. Rather than being confined to the individual that was deemed to possess the capacity, we may discover that the capacity is underwritten by an extended mechanism—a mechanism that reaches beyond the borders/boundaries of the thing to which the capacity is ascribed.[16]

The upshot is that the dispositional hypothesis appears well-equipped to deal with phenomena of the non-cognitive kind.[17] All that is required for the dispositional hypothesis to work is that we have some discernible dispositional property (e.g., a capacity) that, when

[15] As noted by Wilson and Clark (2009), there is no reason why the general notion of an extended process should not be applicable to the realm of computational systems. They suggest that: "computation itself can be an extended process in just the sense in which we are suggesting that cognition can be an extended process" (Wilson & Clark, 2009, p. 60).

[16] There ought to be no reason to doubt the appeal to mechanistic terminology in this scenario. This is despite the fact that we are dealing with a rather unusual set of putative components (e.g., tail flaps, vortices, eddies, and self-generated pressure gradients). Aside from the fact that philosophers have approached the tuna swimming case from a mechanistic standpoint (Kaplan, 2012), practicing scientists seem to have little problem in using mechanistic terminology as part of the effort to explain locomotory phenomena in aquatic settings. The mechanism that explains the tuna's aquatic feats is probably best construed as a propulsive mechanism, and such mechanisms have been the focus of considerable research attention by marine biologists (Fish & Lauder, 2006; Lauder & Drucker, 2002). What is more, a similar appeal to mechanisms can be found in the disciplinary transition to other sorts of hydrological phenomena, such as those to be found in the atmospheric sciences (e.g., Nechayev & Solovyev, 2019). All this is consistent with the purported ubiquity of mechanistic explanations across a multitude of scientific disciplines (Craver & Darden, 2013; Glennan, 2017).

[17] In addition to extended swimming, the dispositional hypothesis can be applied to cases of extended digestion (see Wilson, 2014) and extended respiration (see Di Paolo, 2009, p. 17). In such cases, we credit a biological individual [extended digestion: *Lethocerus*; extended respiration: *Aphelocheirus*] with the possession of a physiological capacity (in this case, a capacity of the digestive/respiratory kind), but the mechanism responsible for the process that reflects the manifestation of this capacity are ones that extend beyond the borders of the entity to which that capacity is ascribed.

manifest, is subject to a form of extended/wide mechanistic realization. This effectively resolves the missing link puzzle, for we can now see how the appeal to non-cognitive kinds (e.g., the physical capacities of various non-human animals) can be accommodated within a broader theoretical framework that also accommodates the notions of extended cognition and the extended mind.

In part the solution to the missing link puzzle stems from the generalizations made in respect of both the dispositional properties (e.g., capacities) that are ascribed to an entity and the constitutive mechanistic phenomena (e.g., cognitive processes) that reflect the exercise of these dispositional properties. In addition to this, however, the resolution of the missing link puzzle stems from the fact that we have generalized the nature of the entities that might be subject to some form of (cognitive or non-cognitive) extension. Together these generalizations provide us with a theoretical account that is applicable to a broad array of disciplines, some of which may lie well beyond the shores of cognitive science.

4.4. Distributed cognition

Despite their differences, the aforementioned cases are all accommodated by the dispositional hypothesis. In one sense, this is a good thing, for we want an account that is sufficiently generic to accommodate cases involving both cognitive and non-cognitive phenomena, as well as cases that fall either side of the extended cognition/extended mind divide. On the other hand, generality is not always a virtue. In particular, we do not want to embrace an account that is overly permissive regarding the kinds of cases that are permitted entry to the club of Extended X.

As a means of addressing this concern, let us attempt to apply the dispositional hypothesis to the notion of distributed cognition (Hutchins, 1995, 2001). Distributed cognition presents us with an interesting challenge, for the nature of the relationship between extended and distributed cognition is not clear-cut. In fact, theorists seldom make an explicit distinction between distributed and extended cognition. Extended cognition is sometimes glossed as a particular form of distributed cognition (e.g., Hutchins, 2011), while, at other times, distributed cognition is seen as a variant of extended cognition (e.g., Carter et al., 2018). Given this, we might expect the dispositional hypothesis to yield a positive response to cases of distributed cognition, thereby confirming the idea that these cognitive kinds are at least closely related.

Interestingly, however, the application of the dispositional hypothesis to distributed cognition yields a negative result. That is, it fails to confirm the status of distributed cognition as a *bona fide* form of extended cognizing. To help us see this, let us direct our attention to what is perhaps the most well-known case of distributed cognition: the case of ship navigation. According to Hutchins (1995), the processes supporting navigational efforts aboard a large maritime vessel are ones that exploit a distributed nexus of biological and non-biological resources. Such resources include multiple human individuals and a rich array of material props, aids, and artifacts. From a mechanistic standpoint, we might say that these resources work together to form a distributed cognitive mechanism that underlies the navigational performances of the ship. The question, of course, is whether this distributed cognitive mechanism ought to be seen as an extended cognitive mechanism, i.e., as a mechanism whose constituents extend beyond the borders of the thing to which the relevant dispositional property (i.e., the navigational ability) is ascribed.

In my view, the answer to this question is "no". The reason for this relates to the difficulty in ascribing ownership of the larger navigational process to one of the components (e.g., a human individual) of the relevant mechanism. In particular, it does not make sense to say that the navigational routine 'belongs' to one of the components of the distributed cognitive mechanism, or that the routine is somehow owned by that component. Nor does it make much sense to say that the relevant navigational ability ought to be seen as a property of one of the individuals (or artifacts) that comprise the larger mechanism. There is, in short, no cognitive 'core' here—some sub-systemic, intra-ship object or agent to which we ascribe a given cognitive ability.[18] Instead, the target of disposition ascription is the ship itself! In determining who or what possesses the relevant ability in the ship navigation case, our attention is naturally drawn to the larger systemic organization (the socio-technical system or ship) that performs the navigational process. We thus say that it is the ship (as a whole) that performs the navigational process, and it is thus the ship (as the whole) that is the target of our ascriptive efforts regarding the possession of specific 'cognitive' abilities, namely, an ability to locate oneself in space or to navigate a course across the ocean waves. The ship navigation case thus fails to qualify as a form of extended cognition. It fails due to the way in which our (disposition-related) ascriptive tendencies are naturally drawn to a larger systemic organization, as opposed to something that counts as a constituent element of that larger organization.

In one sense, this is consistent with the way that Hutchins has attempted to draw a distinction between extended and distributed cognition. Hutchins, for example, suggests that one of the features that distinguishes distributed from extended cognition is that "distributed cognition does not assume a center for any cognitive system" (Hutchins, 2014, p. 37). Note, however, that in earlier work, Hutchins (2011) seeks to downplay the extended theorist's appeal to a cognitive core by referencing work in distributed cognitive science and cultural anthropology. Given the lessons learned from the study of distributed cognitive systems, Hutchins suggests the proponents of extended cognition would be well-advised to eliminate (or at least downplay) the appeal to any sort of cognitive center or cognitive core.

Hutchins is correct, I think, to recognize the role of a cognitive core in distinguishing distributed from extended cognition. But he is wrong to suggest that the proponents of extended cognition ought to dispense with the notion of a cognitive core. The problem is that this core is playing a rather crucial role in arguments for extended cognizing. The core is important, for without this core we have no discernible target for the ascription of cognitive/mental dispositional properties that (when exercised) are subject to extended mechanistic realization. The best we can do, perhaps, is direct our disposition ascriptions to the larger systemic organization that exhibits the disposition manifestation. This is precisely what we do, of course, in the ship navigation case. The problem is that it is hard to see why this ability ought to be regarded as a specifically extended ability, for there is no sense in which the disposition manifestation (the navigational routine) is realized by a mechanism that extends beyond the borders of the thing to which the ability is ascribed. Consider, for example, that if we seek to downplay the status of the human individual as a cognitive core in the long multiplication case, then we are likely to see the multiplicative capacity as a property of the larger system—the system comprising the human individual and pen and paper resources. In an active externalist context, this system would be referred to as an extended cognitive system, but in the absence of a cognitive core, it is nothing more than a distributed cognitive system—a cognitive system whose cognitive capacities (when exercised) are realized by mechanisms that are wholly contained within the borders of the thing to which the capacities are ascribed.

The distinction between distributed and extended cognition thus turns on the way we ascribe cognitive dispositional properties to particular things. In the case of extended cognition, we see a particular entity (e.g., a human individual) as possessing some cognitive ability (e.g., a navigational ability) and the exercise/manifestation of this ability is one that involves the instantiation of an extended mechanism (i.e., a mechanism whose components lie beyond the borders/boundaries of the thing to which the ability is ascribed). This contrasts with the case

[18] See Clark (2008, pp. 106–109), for more on the notion of a cognitive core.

of distributed cognition, where the mechanisms underwriting the ascription of an ability (e.g., a navigational ability) are wholly contained within the borders/boundaries of the thing to which the ability is ascribed (e.g., the ship). These latter mechanisms undoubtedly qualify as *distributed* cognitive mechanisms, in the sense that they are mechanisms that are constituted by an interacting nexus of material objects that include the likes of human individuals and technological artifacts. But there is no sense in which these mechanisms *also* qualify as *extended* cognitive mechanisms.

The upshot is a theoretical distinction between the notions of distributed and extended cognition, one that (for better or worse) challenges the philosophical orthodoxy that has emerged in respect of these cognitive kinds. In contrast to the idea that distributed cognition is a particular form of extended cognition, or that extended cognition is a particular form of distributed cognition, the dispositional hypothesis provides us with an account that distinguishes between these cognitive kinds.[19] This, then, is an additional virtue of the dispositional hypothesis. It tells us something about the distinction between distributed and extended cognition, and it also reveals that efforts to undermine or downplay the notion of a cognitive core (see Hutchins, 2011) are unlikely to advance the cause of the active externalist enterprise. It should also be clear, at this point, that the dispositional hypothesis is not overly liberal regarding the entry of cognitive kinds into the club of Extended X, for it is able to distinguish cognitive kinds that, on the surface at least, look to be highly similar.

4.5. Embedded cognition

While distributed cognition is sometimes confused with extended cognition, philosophers have little problem distinguishing extended cognition from embedded cognition (Rupert, 2004). A key difference between extended and embedded cognition concerns the notions of causal relevance (or dependence) and constitutive relevance (or dependence). In particular, extended cognition is seen to entail a commitment to constitutive relevance, which goes beyond the mere causal relevance insisted upon by the advocates of embedded cognition. As noted by Wheeler (2019b):

> [...] 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, 2019b, p. 861)

The distinction between extended and embedded cognition is easily accommodated by the dispositional hypothesis. The dispositional hypothesis, recall, appeals to the idea of mechanisms whose components lie beyond some border or boundary. Such components are individuated relative to their constitutive relevance to some target phenomenon (see Craver, 2007), which, in the present case, is the constitutive mechanistic phenomenon that reflects the manifestation of a dispositional property (e.g., a cognitive process that reflects the exercise of a cognitive capacity). In this sense, the proponent of the dispositional hypothesis should not be overly concerned about the distinction between embedded and extended cognition. Mechanisms consist of components, and components are individuated courtesy of their constitutive (but not causal) relevance to whatever phenomenon is realized by a mechanism. There is, of course, much to be said about the way in which constitutional claims are to be distinguished from causal claims, both epistemically and metaphysically (see, for example, Baumgartner & Wilutzky, 2017). By itself, however, the nature of this debate does not impugn the status of the dispositional hypothesis. The proponent of the dispositional hypothesis can thus accept that there is more philosophical work to be done in respect of the notion of constitutive relevance, and they can do so without reneging on the basic idea there is something distinctive about constitutional claims—something that distinguishes constitutional claims from merely causal claims.

Perhaps, however, this point about the promissory nature of future philosophical work is the Achilles' heel of the dispositional hypothesis. According to the dispositional hypothesis, we need to be able to identify the components of mechanisms in order to adjudicate the extended status of a mechanism. But if we lack a robust philosophical account of constitutive relevance, then it is somewhat difficult to know how to proceed. If, for example, we cannot be sure that an extra-organismic resource is a *bona fide* component of a putatively extended mechanism, then we have no means of applying the dispositional hypothesis to candidate cases of cognitive extension. In this case, the practical value of the dispositional hypothesis is, at best, limited.

In response to this, it is worth bearing in mind that scientists seem to have little problem in individuating the components of the mechanisms. This is not to say that scientists do not need to invest considerable effort in individuating such components; it is simply to say that they do, as a matter of fact, discover mechanisms, and such mechanisms consist of a causally interacting nexus of what (in neo-mechanical philosophy) are referred to as components. From a practical standpoint, then, it is hard to see how the absence of a philosophically-robust account of constitutive relevance could stymie the scientific effort to discover and describe extended mechanisms.

The project of individuating components is, if anything, even easier when it comes to engineering disciplines, for engineers already know a great deal about the componential structure of the mechanisms that they themselves create.[20] This, of course, should come as no great surprise, for engineers design mechanisms to perform certain functions, and the constituents (i.e., components) of these mechanisms are specifically selected so as to achieve the desired functionality.[21] (The relevance of this should be clear when it comes to the effort to design and build extended cognitive systems.)

From a purely practical standpoint, then, there seems little reason to doubt the importance of the distinction between constitutional and causal claims. Nor does there seem any reason to think that the *practical* project of studying and (crucially) building extended cognitive systems is apt to be stymied by the absence of a philosophically-robust account of constitutive relevance. To be sure, it would certainly help to have such an account to hand.[22] But there is no reason why the dispositional hypothesis cannot be used to support the practical pursuits of

[19] Much depends, of course, on how we define the notion of distributed cognition. If, for example, we see distributed cognition as nothing more than a claim about the distributed nature of cognitive mechanisms—a characterization that applies to cognitive mechanisms of any stripe (extended or otherwise), then it should be clear that extended cognition will emerge as a particular form of distributed cognition, as will non-extended (e.g., brain-based) forms of cognition. This appears to be the view endorsed by Hutchins (2014, p. 36), who suggests that "Distributed cognition is not a kind of cognition; it is a perspective on all of cognition... Distributed cognition begins with the assumption that all instances of cognition can be seen as emerging from distributed processes".

[20] As noted by Wilson and Clark (2009, p. 63), "An electronics engineer usually has a pretty clear sense of what is mere input to a system and what is an integrated addition that alters the system itself".

[21] Such mechanisms are what Glennan and Illari (2018b) refer to as mechanisms with designed-and-built-etiologies.

[22] For a recent philosophical account of constitutive relevance, see Craver et al. (2021).

(extended) cognitive science. Nor is there any reason to think that such pursuits are the peculiar province of cognitive science, for the dispositional hypothesis is one that can be applied to multiple kinds of phenomena, not just those that pique the interests of the cognitive scientific community.

5. Conclusion

In the present paper, I sought to provide a theoretical account that extends the reach of active externalism to the realm of both cognitive and non-cognitive phenomena (the realm of what I called Extended X). This hypothesis – dubbed the dispositional hypothesis – assigns a central role to dispositional properties, where the notion of a dispositional property subsumes the likes of capacities, abilities, and dispositional beliefs. According to the dispositional hypothesis, extended cognition occurs when the mechanism responsible for the manifestation of a (cognitive) dispositional property includes components that lie external to the borders/boundaries of the thing to which the dispositional property is ascribed.

One of the immediate virtues of the dispositional hypothesis is that it provides us with a common approach to understanding extended cognition and the extended mind—the two most prominent targets of the active externalist endeavor. While we might be inclined to view active externalism as something of a dualistic enterprise, with the relevant philosophical space partitioned into regions concerned with either the extended mind (the philosophy of mind) or extended cognition (the philosophy of cognitive science), the dispositional hypothesis papers over these distinctions. According to the dispositional hypothesis, we can understand claims about the extended mind in precisely the same way we understand claims about extended cognition; all that changes is the nature of the dispositional property that is ascribed to a given cognitive agent.

A second virtue of the dispositional hypothesis is the way it broadens the scope of active externalism. While active externalism is mostly concerned with the realm of cognitive/mental phenomena, the dispositional hypothesis opens the door to a more liberal view of active externalism—one that supports its application to phenomena that lie beyond the disciplinary borders of cognitive science.

A third virtue of the dispositional hypothesis is the way it helps us make sense of the appeal to mechanistic concepts in the active externalist literature. The dispositional hypothesis is thus consistent with the idea that cognitive extension can be understood from a broadly mechanistic perspective (e.g., Fazekas, 2013). It also provides us with a particular way of understanding the appeal to extended (Clark, 2011; Hurley, 2010; Kaplan, 2012; Smart, 2022; Zednik, 2011), wide (Miłkowski et al., 2018), or supersized (Clark, 2008) mechanisms in the active externalist literature.

Is the dispositional hypothesis the right way to think about active externalism? The dispositional hypothesis is, I think, extensionally adequate, in the sense that it is applicable to a diverse array of Extended X phenomena. That being said, there is clearly much more work to be done when it comes to the philosophical analysis of dispositional properties, the link between dispositional properties and mechanistic concepts, and the extent to which the dispositional hypothesis successfully discriminates between dispositions of the extended versus non-extended kind. These are all important areas for future philosophical research. For present purposes, however, I hope to have shown that the dispositional hypothesis provides us with a plausible approach to understanding cognitive extension, one that accommodates many of the cases discussed in the active externalist literature. It provides us with an explicit characterization of what might be entailed by active externalist claims, and it serves as an important alternative to recent philosophical accounts that emphasize the importance of sensorimotor exchanges between a cognitive agent and the extra-agential environment (see Chalmers, 2019). It also provides us with a means of linking active externalist debates and discussions to phenomena that lie beyond the orbit of cognitive science. In this respect, the dispositional hypothesis provides us with an important opportunity to broaden the scope of active externalist theorizing, enabling us to extend its reach beyond the borders of cognitive science and the philosophy of mind. As a philosophical theory, active externalism was forged in a *cognitive* scientific crucible, and it has transformed the way we think about the machinery of the *human* mind. Perhaps, however, the human cognitive crucible was just the beginning. By broadening the scope of active externalism to the realm of Extended X, we may shed light on phenomena that lie beyond the borders cognitive science. And it is perhaps via that circuitous loop out into the extra-cognitive realm that we may arrive at a better understanding of just what it means for the human mind to escape its cranial confines and seep out into the world.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Adams, F. (2010). Why we still need a mark of the cognitive. *Cognitive Systems Research*, *11*(4), 324–331.

Adams, F., & Garrison, R. (2013). The mark of the cognitive. *Minds and Machines*, *23*(3), 339–352.

Allen-Hermanson, S. (2013). Superdupersizing the mind: Extended cognition and the persistence of cognitive bloat. *Philosophical Studies*, *164*(3), 791–806.

Baumgartner, M., Casini, L., & Krickel, B. (2020). Horizontal surgicality and mechanistic constitution. *Erkenntnis*, *85*, 417–340.

Baumgartner, M., & Wilutzky, W. (2017). Is it possible to experimentally determine the extension of cognition? *Philosophical Psychology*, *30*(8), 1104–1125.

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.

Carter, J. A., Clark, A., & Palermos, S. O. (2018). 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, UK: Oxford University Press.

Carter, J. A., & Kallestrup, J. (2016). Extended cognition and propositional memory. *Philosophy and Phenomenological Research*, *92*(3), 691–714.

Cartwright, N. (2007). What makes a capacity a disposition? In M. Kistler, & B. Gnassounou (Eds.), *Dispositions and causal powers* (pp. 195–205). Aldershot, UK: Ashgate Publishing Limited.

Chalmers, D. J. (2019). Extended cognition and extended consciousness. In M. Colombo, E. Irvine, & M. Stapleton (Eds.), *Andy clark and his critics* (pp. 9–20). New York, New York, USA: Oxford University Press.

Clark, A. (1997). *Being there: Putting brain, body and world together again*. Cambridge, Massachusetts, USA: MIT Press.

Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. New York, New York, USA: Oxford University Press.

Clark, A. (2011). Finding the mind. *Philosophical Studies*, *152*(3), 447–461.

Clark, A. (2015). What 'Extended Me' knows. *Synthese*, *192*(11), 3757–3775.

Clark, A. (2019). Replies to critics: In search of the embodied, extended, enactive, predictive (EEE- P) mind. In M. Colombo, E. Irvine, & M. Stapleton (Eds.), *Andy clark and his critics* (pp. 266–302). New York, New York, USA: Oxford University Press.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, *58*(1), 7–19.

Craver, C. (2007). Constitutive explanatory relevance. *Journal of Philosophical Research*, *32*, 3–20.

Craver, C. F., & Darden, L. (2013). *In search of mechanisms: Discoveries across the life sciences*. Chicago, Illinois, USA: The University of Chicago Press.

Craver, C. F., Glennan, S., & Povich, M. (2021). Constitutive relevance & mutual manipulability revisited. *Synthese*, *199*(3), 8807–8828.

Cummins, R. (2000). "How does it work?" versus "what are the laws?": Two conceptions of psychological explanation. In F. C. Keil, & R. A. Wilson (Eds.), *Explanation and cognition* (pp. 117–144). Cambridge, Massachusetts, USA: MIT Press.

Di Paolo, E. (2009). Extended life. *Topoi*, *28*(1), 9–21.

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.

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). Kirchberg am Wechsel, Austria: Austrian Ludwig Wittgenstein Society.

Fish, F. E., & Lauder, G. V. (2006). Passive and active flow control by swimming fishes and mammals. *Annual Review of Fluid Mechanics*, *38*, 193–224.

Glennan, S. (2017). *The new mechanical philosophy*. Oxford, UK: Oxford University Press.
Glennan, S., & Illari, P. M. (Eds.), (2018). *The Routledge handbook of mechanisms and mechanical philosophy*. New York, New York, USA: Routledge.
Glennan, S., & Illari, P. M. (2018). Varieties of mechanisms. In S. Glennan, & P. M. Illari (Eds.), *The Routledge handbook of mechanisms and mechanical philosophy* (pp. 91–103). New York, New York, USA: Routledge.
Huneman, P. (2013). Causal parity and externalisms: Extensions in life and mind. *Minds and Machines*, *23*(3), 377–404.
Hurley, S. (2010). The varieties of externalism. In R. Menary (Ed.), *The extended mind* (pp. 101–153). Cambridge, Massachusetts, USA: MIT Press.
Hutchins, E. (1995). *Cognition in the wild*. Cambridge, Massachusetts, USA: MIT Press.
Hutchins, E. (2001). Distributed cognition. In N. J. Smelser, & P. B. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (pp. 2068–2072). Oxford, UK: Elsevier Science Ltd.
Hutchins, E. (2011). Enculturating the supersized mind. *Philosophical Studies*, *152*(3), 437–446.
Hutchins, E. (2014). The cultural ecosystem of human cognition. *Philosophical Psychology*, *27*(1), 34–49.
Hüttemann, A., & Kaiser, M. I. (2018). Potentiality in biology. In K. Engelhard, & M. Quante (Eds.), *Handbook of potentiality* (pp. 401–428). Dordrecht, The Netherlands: Springer.
Japyassú, H. F., & Laland, K. N. (2017). Extended spider cognition. *Animal Cognition*, *20*(3), 375–395.
Jonker, C. M. (2008). Extended mind from an artificial intelligence perspective. In L. Zonneveld, H. Dijstelbloem, & D. Ringoir (Eds.), *Reshaping the human condition: exploring human enhancement* (pp. 133–140). The Hague, Netherlands: Rathenau Institute.
Kaiser, M. I., & Krickel, B. (2017). The metaphysics of constitutive mechanistic phenomena. *The British Journal for the Philosophy of Science*, *68*(3), 745–779.
Kaplan, D. M. (2012). How to demarcate the boundaries of cognition. *Biology & Philosophy*, *27*(4), 545–570.
Kirsh, D. (2009). Problem solving and situated cognition. In P. Robbins, & M. Aydede (Eds.), *The Cambridge handbook of situated cognition* (pp. 264–306). Cambridge, UK: Cambridge University Press.
Lauder, G. V., & Drucker, E. G. (2002). Forces, fishes, and fluids: hydrodynamic mechanisms of aquatic locomotion. *Physiology*, *17*(6), 235–240.
Law, E., & von Ahn, L. (2011). Human computation. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, *5*(3), 1–121.
McKenna, R. (2019). Extended epistemology. *Analysis*, *79*(4), 790–799.
Menary, R. (Ed.), (2010). *The extended mind*. Cambridge, Massachusetts, USA: MIT Press.
Menary, R. (2015). Mathematical cognition: A case of enculturation. In T. K. Metzinger, & J. M. Windt (Eds.), *Open MIND: Philosophy and the mind sciences in the 21st century* (pp. 1–20). Frankfurt am Main, Germany: MIND Group.
Miłkowski, M., Clowes, R., Rucińska, Z., Przegalińska, A., Zawidzki, T., Krueger, J., Gies, A., McGann, M., Afeltowicz, L., Wachowski, W., Stjernberg, F., Loughlin, V., & Hohol, M. (2018). From wide cognition to mechanisms. *Frontiers in Psychology*, *9*(Article 2393), 1–17.
Mumford, S. (1998). *Dispositions*. Oxford, UK: Oxford University Press.
Mumford, S. (2009). Causal powers and capacities. In H. Beebee, C. Hitchcock, & P. Menzies (Eds.), *The Oxford handbook of causation* (pp. 265–278). Oxford, UK: Oxford University Press.
Nechayev, A. M., & Solovyev, A. (2019). On the mechanism of atmospheric vortex formation and how to weaken a tornado. *European Journal of Applied Physics*, *1*(1), 1–5.
Palermos, S. O. (2014). Loops, constitution, and cognitive extension. *Cognitive Systems Research*, *27*, 25–41.
Parise, A. G., Gagliano, M., & Souza, G. M. (2020). Extended cognition in plants: is it possible? *Plant Signaling & Behavior*, *15*(2), Article 1710661.
Pöyhönen, S. (2014). Explanatory power of extended cognition. *Philosophical Psychology*, *27*(5), 735–759.
Roberts, T. (2012). You do the maths: Rules, extension, and cognitive responsibility. *Philosophical Explorations*, *15*(2), 133–145.
Rupert, R. D. (2004). Challenges to the hypothesis of extended cognition. *Journal of Philosophy*, *101*(8), 389–428.
Sims, M., & Kiverstein, J. (2022). Externalized memory in slime mould and the extended (non-neuronal) mind. *Cognitive Systems Research*, *73*, 26–35.
Smart, P. R. (2018). Human-extended machine cognition. *Cognitive Systems Research*, *49*, 9–23.
Smart, P. R. (2022). Toward a mechanistic account of extended cognition. *Philosophical Psychology*, *35*(8), 1107–1135.
Triantafyllou, M. S., & Triantafyllou, G. S. (1995). An efficient swimming machine. *Scientific American*, *272*(3), 64–71.
van Gelder, T. (1995). What might cognition be, if not computation? *The Journal of Philosophy*, *92*(7), 345–381.
Wheeler, M. (2010). In defense of extended functionalism. In R. Menary (Ed.), *The extended mind* (pp. 245–270). Cambridge, Massachusetts, USA: MIT Press.
Wheeler, M. (2019). Breaking the waves: Beyond parity and complementarity in the arguments for extended cognition. In M. Colombo, E. Irvine, & M. Stapleton (Eds.), *Andy clark and his critics* (pp. 81–95). New York, New York, USA: Oxford University Press.
Wheeler, M. (2019). The reappearing tool: transparency, smart technology, and the extended mind. *AI & Society*, *34*(4), 857–866.
Wilson, R. A. (2010). Extended vision. In N. Gangopadhyay, M. Madary, & F. Spicer (Eds.), *Perception, action and consciousness: Sensorimotor dynamics and two visual systems* (pp. 277–290). Oxford, UK: Oxford University Press.
Wilson, R. A. (2014). Ten questions concerning extended cognition. *Philosophical Psychology*, *27*(1), 19–33.
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, UK: Cambridge University Press.
Wilson, R. A., & Craver, C. (2007). Realization: Metaphysical and scientific perspectives. In P. Thagard (Ed.), *Philosophy of psychology and cognitive science* (pp. 81–104). Oxford, UK: North-Holland.
Zednik, C. (2011). The nature of dynamical explanation. *Philosophy of Science*, *78*(2), 238–263.