Juraj Hvorecký
Institute of Philosophy, Czech Academy of Sciences

Marcin Miłkowski
Institute of Philosophy and Sociology, Polish Academy of Sciences

Theoretical virtues of cognitive extension

Abstract

This paper argues that the extended mind approach to cognition can be distinguished from its alternatives, such as embedded cognition and distributed cognition, not only in terms of metaphysics, but also in terms of epistemology. In other words, it cannot be understood in terms of a mere verbal redefinition of cognitive processing. This is because the extended mind approach differs in its theoretical virtues compared to competing approaches to cognition. The extended mind approach is thus evaluated in terms of its theoretical virtues, both essential to empirical adequacy and those that are ideal desiderata for scientific theories. While the extended mind approach may have similar internal consistency and empirical adequacy compared to other approaches, it may be more problematic in terms of its generality and simplicity as well as unificatory properties due to the cognitive bloat and the motley crew objections.

1. Introduction

The goal of this paper is to argue that the extended mind approach can be distinguished from its alternatives not only in terms of metaphysics, but also epistemology. In other words, we claim that this view differs in its theoretical virtues compared to other similar approaches to cognition.

To accomplish this, we will proceed as follows: In Section 2, we will introduce the objection against the extended mind view on cognition that states that it does not differ from its competitors in terms of the empirical content of the explanations involved. We will use the example of the audience effect as framed in terms of extended emotional processing to illustrate this point. Then, we will argue that there are further criteria of theory choice, traditionally understood in terms of theoretical virtues, that are implied by the extended view and its alternatives. Specifically, we will argue that previous discussions concerning the advantages and disadvantages of the approach lead to the observation that it may have problematic consequences for the generality and simplicity as well as unificatory properties of cognitive theories. In this regard, the extended view, particularly in its more recent versions, has specific epistemic implications for the development of cognitive theories. At the same time, we will note that the theoretical virtues of general approaches to cognition (or theoretical frameworks) should not be confused with those of particular explanations of phenomena. In conclusion, we will state that the extended view is epistemically distinguishable from its competitors. Therefore, the objection can be rebutted.

However, our conclusion is not particularly optimistic for the view, as the difference does not seem to be in its favor due to considerations of generality, parsimony, and unification. While it is certainly possible for defenders of the extended cognition to develop their theoretical framework to warrant a more systematic approach to developing explanations that remain appropriately general, parsimonious, and unificatory, the way this generality, parsimony, and unification could be achieved will most likely be quite different from traditional approaches that rely on invariant structures in individuals or cognitive systems to warrant these desiderata.

2. The issue of empirical equivalence

In this section, we will address the issue of the empirical equivalence of the extended mind approach with similar, but metaphysically distinct frameworks in contemporary cognitive science. Essentially, if the extended mind view is indistinguishable from alternative views in terms of empirical evidence, then the utility of engaging in the metaphysics of the extended mind could be called into question and seen as a futile exercise (Müller, 2018). Additionally, some have argued that there is no empirical difference between the situated cognition approach and the extended mind approach (Barker, 2010; Sprevak, 2010). Barker states that differences between the two approaches are not empirically testable and that they cannot be distinguished based on theoretical virtues, in particular simplicity and unification. In short, they are equivalent in terms of their empirical accuracy, testability, and other theoretical virtues. Sprevak argues that by looking at the cognitive science practice, one cannot actually settle the debate between the two approaches, as cognitive science is not sensitive to the difference between causal and constitutive relevance. If these arguments hold, then it becomes unclear what, if any, relevance the extended mind view holds for empirical research. This also undermines the original claims of Clark and Chalmers, who stated that their view has “obvious consequences for philosophical views of the mind and for the methodology of research in cognitive science” (Clark & Chalmers, 1998, p. 18).

Before we can assess this claim, it is important to clarify how we understand the terms “situated cognition”, “extended cognition”[1], and “distributed cognition”. These terms all refer to approaches that fall under the umbrella of 4E cognition (embodied, embedded, extended, and enactive), or wide cognition (Miłkowski et al., 2018). They are also sometimes linked to the theory of niche construction, which posits that an organism's choices, activities, and metabolic processes shape or modify the environment (Laland et al., 2000; for a more detailed examination in the cognitive and emotional domain, see Krueger, 2014). The situated approach to cognition, also known as "embedded cognition," holds that cognition should be understood in terms of the (typically time-sensitive) interaction between the agent and its immediate surroundings. According to it, the extra-bodily context constrains and enables cognition. The extended mind idea suggests that cognitive processes are not necessarily limited to the brain and can incorporate external resources such as tools, language, and external systems in the environment. Resulting cognitive processes literally incorporate the external elements as their integral parts. Finally, the distributed cognition approach does not focus on cognition as a property of an individual organism or agent. Rather, it describes

[1] In this paper, we use the terms "extended mind" and "extended cognition" interchangeably.

larger cognitive systems, which may encompass multiple individual agents and artifacts. Within the distributed framework, cognitive states often result from interactions among several agents and are shared among them.

At first glance, these approaches may appear to differ in their claims. However, upon closer examination, it may be that there is no objective way to distinguish their explanatory achievements. For example, let's consider an agent A, who employs a tool T_1 in their cognitive process. According to situated cognition, T_1 is considered as part of the environment, whereas in extended cognition, it is regarded as part of A's mind. Under the framework of distributed cognition, T_1 is seen as part of a larger cognitive system that encompasses A and T_1. Nonetheless, the causal impact of T_1 remains the same across all three cases. Within all three frameworks, the attainment of a specific goal would not have been possible without the interaction between A and T_1. In other words, the causal structure responsible for the cognitive process remains the same in all three cases, with the only difference lying in the labels attributed to A and T_1. At best, these labels could be seen as interpretive glosses, but their theoretical significance is questionable as their explanations proceed in the same manner. Thus, the argument can be presented as follows:

1. Extended, situated, and distributed explanations of a cognitive phenomenon C posit the same causal structures responsible for C.
2. If an explanation E_1 posits the same causal structure for C as E_2, E_1 is explanatorily (epistemologically) equivalent to E_2.

Hence, extended, situated, and distributed explanations of C are explanatorily equivalent.

This is an impeccable *modus ponens*. While in actual cases, the specific approaches may differ in the proposed causal structures, this is only because researchers tend to make additional causal assumptions. In reality, all situated cognition explanations could be rephrased in terms of extended or distributed cognition and *vice versa* without any loss of empirical meaning. This implies that the entire debate over extended mind or distributed cognition is purely metaphysical. This may put supporters of wide cognition in a difficult position, as it suggests that the extended mind thesis is indeed purely a matter of terminology (Müller, 2018).

However, there is a problem with the argument presented above. The issue is that causal structures are always posited to explain a particular phenomenon, and how one defines the phenomenon to be explained is always influenced by one's theoretical perspective (Craver, 2009). All three approaches, if fact, imply different theoretical perspectives on cognitive phenomena. This means that their definitions of the phenomena to be explained will inevitably differ, thus the causal structures must be different as well. We can illustrate this with the example of the audience effect, which can be understood through the lens of extended emotional processing (Krueger, 2014; Krueger & Szanto, 2016).

The “audience effect” (Fernandez-Dols & Ruiz-Belda, 1997) is grounded in the observation that individuals exhibit more pronounced emotional expressions in the presence of others compared to when they are alone. This effect is particularly evident in sport situations where strong emotions and crowds are both present. For example, it has been observed that upon a successful strike, bowling players tend to smile more intensely when turning towards their fellow players, rather than when hitting the pins (Kraut & Johnston, 1979). This is also the case in what may be considered the pinnacle of a sporting career, such as winning at the Olympics. During award ceremonies, athletes tend to express the most joy during the handshake with an Olympic official, when social interaction is most pronounced. They are significantly less likely to express their joy when alone, without the presence of a perceptive crowd. In this regard, it would be interesting to compare the expressions of joy of athletes during regular award ceremonies to similar situations without crowds (such as the Tokyo Olympics of 2022 due to the COVID-19 situation). The audience effect, which requires external stimulation, stands in sharp contrast to internalist theories of emotions that emphasize the role of mandatory emotional expression stemming solely from within the organism (Ekman et al., 2013). Ekman and his collaborators take emotion expressions to be externally directed demonstrations of inner processes that take place within the individual, fully in line with internalist principles.

The key finding from studies on the audience effect is that the intensity of emotional reactions, as seen in facial expressions, is influenced by social interactions and reinforced by the audience to whom it is directed. However, it is not a one-way projection from the athlete to the crowd. Instead, the emotion is evoked by the presence of the audience, making the emotion expression stronger for the athlete. The emotions under discussion are a result of the coupling between individual athletes and onlookers. This does not mean that people do not experience emotional episodes when alone. The effect only shows that the intensity of emotion expression varies with the presence of involved observers. There are even suggestions that in solitary scenarios, implicit sociality in the form of imagined or hypothesized observers can enhance emotional reaction (Fridlund, 1991).

The audience effect serves as a prime example of a mental phenomenon that surpasses the boundaries of the individual. All three frameworks of embedded, extended, and distributed cognition have the capability to account for this phenomenon, albeit from different perspectives. The embedded mind framework directs its attention towards the situated nature of the individual mind within a specific environment, such as a stadium, to comprehend the effect. On the other hand, the extended mind approach proposes that the expression of emotions is co-constructed by the cheering audience, in addition to the internal input from the athlete. On this view, members of the crowd are actively involved in the experience of joyful emotion, contributing to its sustenance. Meanwhile, the perspective of distributed cognition posits that the mental state of triumphant joy is both located and shared among all participating actors, including the athletes, the crowd, and other relevant stakeholders.

As we've seen, the theoretical posits each of the three frameworks bring in differ

substantially, but these differences do not prevent each framework from explaining the audience effect phenomenon successfully. In fact, even from the brief description of their explanatory strategies, it becomes clear that the main difference remains largely metaphysical. Each approach introduces a different ontology for mental entities and offers distinct conditions for their individuation and persistence. However, Barker (2010) argued that these differences were not testable. The varying interpretations of the effect are yet another reason to express skepticism about the precise ontology of mental states, as well as the futility of the concurring frameworks. Everyone agrees that the phenomenon cannot be explained without reference to external events. Thus, the precise understanding of the nature of mental states is, at best, secondary.

Before we proceed, an important caveat is in order. Extended or embodied cognition, as frameworks, possess limited, if any, inherent explanatory power. However, this is not fundamentally different from other types of frameworks, such as computationalism (Miłkowski, 2018; Wołoszyn & Hohol, 2017). The reason is that, as frameworks, their primary role is not to directly explain specific phenomena. Instead, they serve as inspiration and guidance for researchers in constructing general explanatory theories or specific models for the phenomena under investigation. As we will delve into below, this caveat isn't a mere terminological quibble; there are deeper nuances to explore.

In this section, we have outlined the problem of the empirical equivalence of the extended mind approach with similar, but metaphysically distinct frameworks in contemporary cognitive science. We have shown that despite different theoretical posits, all three frameworks can successfully explain the audience effect phenomenon. However, this leaves us with a purely metaphysical debate, with no clear way to test the ontological differences between the frameworks. While some may argue that this ends the discussion, we argue that by examining theoretical virtues, there may be reasons to prefer one version of the wide approach over others, *contra* Barker. In the next section, we will delve deeper into this examination.

3. Problems with generality

In this section, we argue that even though the extended mind approach and its alternatives may exhibit explanatory equivalence, there are still reasons to prefer certain versions of a particular approach over others. To make this case, we will examine the theoretical virtues of these approaches. These virtues include accuracy, breadth of scope (or generality), consistency, simplicity, and fruitfulness, as outlined by philosophers of science such as Kuhn (1977). While for space reasons, we won't be able to examine all of these virtues in depth, we will provide a brief overview of each of them to set the stage for their later use in our analysis. Subsequently, we will focus on generality and unification.

Accuracy relates to a theory’s ability to adequately account for specific phenomena and its alignment with various experimental results. Consistency can be evaluated in both internal and external forms. Internal consistency assesses whether a theory's claims are

non-contradictory, while external consistency evaluates its compatibility with other accepted theories within the same field and beyond. The breadth of scope refers to the range of phenomena that a theory addresses. Simplicity not only relates to the number of entities introduced and the manageability of the theoretical apparatus, but also to the assumption that the theory introduces order among previously divergent phenomena. Finally, fruitfulness refers to a theory's ability to discover new phenomena and/or relationships between already known ones, as well as its potential for further productivity. This criterion is synonymous with a theory's ability to initiate or sustain a viable research program (Lakatos, 1970) or tradition (Laudan, 1977).

We are aware that in the course of his thinking on the strength and utility of the five criteria, Kuhn expresses skepticism about an unproblematic way to make simple comparisons between competing theories (see, for example, Okasha, 2011). Furthermore, Kuhn's list is not definitive, and in the course of further discussions, longer and more structured lists have been produced (for a recent one, see Keas, 2018). We do not claim that a particular list is complete, nor that there might not be disputes about how to interpret each criterion. Our argument is that there are some *epistemological* criteria for theory choice that go beyond explanatory adequacy and purely metaphysical debates, and they might prove useful when employed in the evaluation of competing approaches. For us, Kuhn's original list is a starting point for a debate that has implications for comparisons between situated, extended, and distributed mind frameworks. As we aim to show, on some criteria all frameworks come out the same, but there are dimensions in theory choice where substantial differences appear, and some approaches will score low on them, which might have consequences for their broader acceptability.

In the remainder of this section, we will evaluate the three frameworks of situated, extended, and distributed cognition in terms of the theoretical virtues as analyzed by Douglas (2013). Douglas divides these virtues into four groups: the first two groups include minimal criteria for adequate science, and the other groups are desiderata for theories per se, and desiderata for theories in relation to evidence. Group 1 includes internal consistency (but see (Vickers, 2013) for an argument that science can proceed with inconsistent theories), and group 2 involves empirical adequacy. Group 3 covers desiderata when applied to theories *per se*, such as scope, simplicity, and explanatory power, while group 4 includes desiderata when applied to theories in relation to evidence, such as unification (understood in terms of explanatory scope, simplicity, external consistency and coherence), novel prediction, and precision (see Figure 1 for a summary). While generality may be assessed without knowing any evidence, unification cannot: “generality” refers to the ability of a theory to apply to a wide range of phenomena, while unification refers to the ability of a theory to bring together diverse phenomena under a single explanatory framework, and these phenomena must be actually accounted for, which requires us to know the evidence.

We assume, in line with Bernecker (2014), that all three approaches and their associated explanations do not differ with respect to minimal criteria. In particular, if the extended cognition approach is empirically equivalent to the other two approaches, then it

must be meeting the criteria from group 2, such as internal consistency and empirical adequacy. The differences appear in the ideal desiderata, particularly in group 3, which include factors such as scope, simplicity, and explanatory power, and by implication, in group 4 (as unification depends on the scope, simplicity.)

We will argue that the issue of cognitive bloat, frequently mentioned in criticisms of extended cognition, is related to virtues in group 3. This is because cognitive bloat is supposed to be detrimental to simplicity and generality (or scope). Similar considerations are at play in the "motley crew argument" against wide approaches to cognition (Casper, 2023; Shapiro, 2011, p. 198).

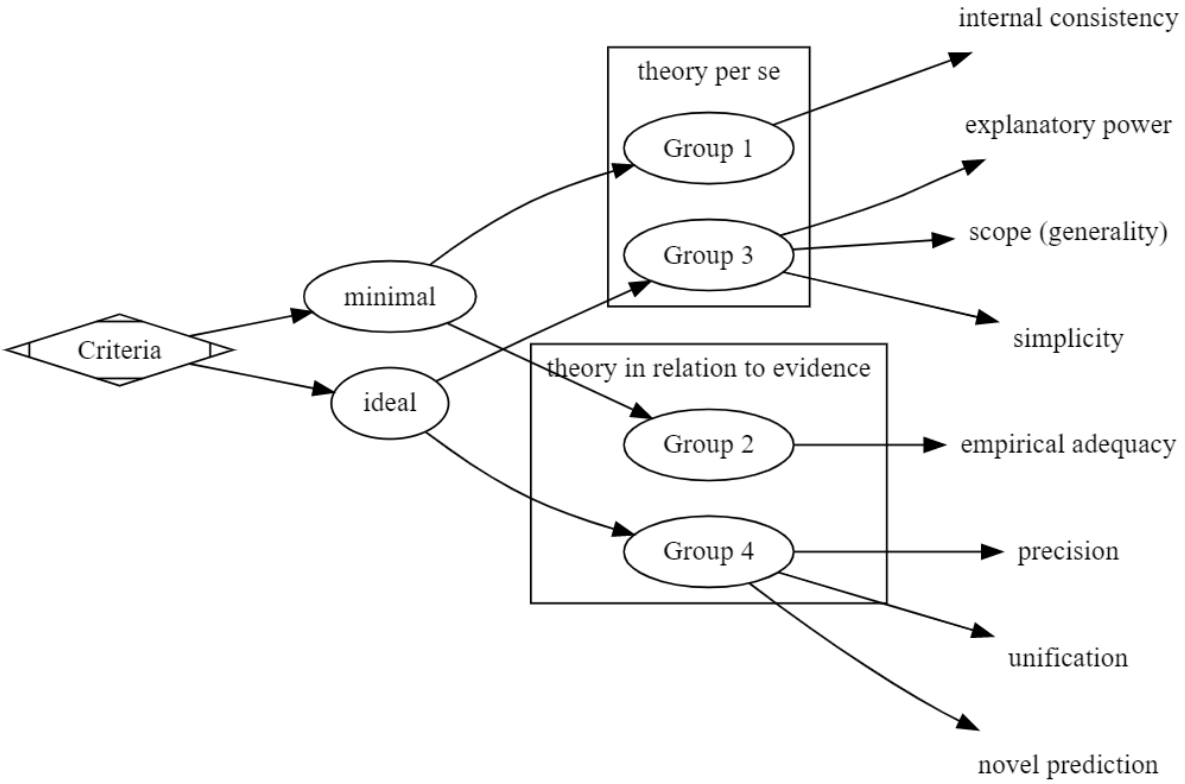

Figure 1: Criteria of theory choice according to Douglas (2013), as grouped by their significance (minimal or ideal), and by their target (theory per se or in relation to evidence). Diagram by the authors.

As we have indicated, the fact that these frameworks are not equivalent is borne out by analyzing the controversy surrounding the cognitive bloat and the motley crew argument. According to Rowlands, the cognitive bloat objection rests on the observation that:

> the admission of extended cognitive processes places one on a slippery slope. Once we permit such processes, where do we stop? Our conception of the cognitive will become too permissive, and we will be forced to admit into the category of the

> cognitive all sort sorts [sic!] of structures and processes that clearly are not cognitive. (Rowlands, 2009, p. 2)

At the level of individual explanations, the criterion of generality, or the breadth of scope (Thagard speaks of consilience in his 1978), seems central in this context. It refers to the ability of an explanation to accommodate numerous instances, as well as generalize to newer occurrences. Commitments of various frameworks put forward distinct restrictions to their ability to generalize.

Generality in cognitive theories has long concerned cognitive scientists, as they have been wary of the potential fragmentation that could impede informative generalizations. An effective method to attain this generality is by grounding theories in invariant structures. For instance, Newell and Simon argued that only a few fundamental characteristics of information processing systems, which underlie human problem-solving, remain invariant across tasks and problem solvers (Newell & Simon, 1972, p. 788). To enhance the invariance of the system's structure, Simon even went to the extent of excluding memory as a component, instead referring to it as the internal environment. This is due to the fact that memory is not invariant over time and undergoes changes during the learning process, rendering it challenging to derive generalizations about the architecture of the cognitive system. Rather than expanding the cognitive system to support generalizations, Newell and Simon sought to minimize the extent of its structure.

As we understand it, the "cognitive bloat" objection is not just about the extent to which the notion of cognition is extended, but also about the stability of this extension over time. The concern is that if the system "bloats" to include parts that are not relevant to other cognitive tasks, then both the generality and simplicity of the cognitive system will be at risk. This is because the simple, invariant structure of the cognitive system can no longer be assumed to exist.

In contrast to the extended cognition framework, the situated cognition approach posits that individual mental states are shaped by the environment and the cognitive architecture of the individual remains constant. Proponents of this approach argue that biological individuals are stable entities that can be used to explain a variety of cognitive phenomena (Rupert, 2009). As a result, the situated cognition approach has the potential for a broader scope while still being parsimonious, since the invariant structure remains the same across multiple explanations of various phenomena.

Distributed cognition is also less vulnerable to the cognitive bloat objection. This is because the scope of the cognitive system, which can involve multiple individuals and artifacts, is not necessarily required to extend into the environment at all. In principle, this approach can accommodate numerous phenomena that rely solely on the capacities of a single biological individual. In contrast, the extended mind approach aims to demonstrate that the mind does indeed extend in most cases. Consider a simple comparative scenario: a human solving a mnemonic task by memorizing a short poem. For proponents of cognitive

extension, language serves as a prime example of the mind's extension. However, it is worth noting that, in relation to the aforementioned cognitive bloat issue, advocates of the extended framework owe us a more precise explanation regarding which aspect of language extends the mind of an individual learning the poem. There appears to be no principled way to ascertain whether the individual mind has been extended solely by the words of the poem, the entire vocabulary of language, the rules for sentence formation, abstract linguistic forms, or some other factor altogether. This problem is not as pressing for the proponent of distributed cognition since, within this framework, a poem is simply considered a physical artifact fulfilling a cognitive role, and it does not necessarily need to be included as part of an individual's mind in order to explain their success in remembering it.

What may seem even more surprising is that the cognitive bloat problem can be considered more grave in some recent renditions of the extended mind approach, which has been understood to come in at least three waves (Sutton, 2010). The first version, or wave, of the extended mind thesis rested on the argument that artifacts were functionally equivalent to parts of the mind, which was the core of the parity argument in (Clark & Chalmers, 1998). The second wave sees artifacts as complementing (and not necessarily analogous to) the mind. For example, mathematical notations need not be analogous or isomorphic to the mind's language of thought to be considered prime examples of cognitive extension (Menary, 2007, 2015). Finally, the third wave sees artifacts and sociomaterial culture in general as transforming the mind, rather than merely complementing it (Kirchhoff & Kiverstein, 2019). The bloat objection has been posed against the first wave (Allen-Hermanson, 2013), but it applies to making artifacts parts of the mind, and thereby analogously to the second wave as well. It is, however, the most critical for the third wave: if the individuals are "dissolved into peculiar loci of coordination and coalescence among multiple structured media" (Sutton, 2010, p. 213), then they do not seem poised to be invariants that support important theoretical generalizations. In fact, it becomes an issue whether the third wave leaves space for any invariants at all.

The motley crew argument points to a similar difficulty: it states that "processes that cross the bounds of the brain are not well-defined – are a motley crew – and so cannot be an object of scientific investigation" (Shapiro, 2011, p. 198). Shapiro notes that this should be understood as implying that transcranial processes are not a "well-formed kind" (p. 199). Artifacts we use differ from natural kinds in that they do not support lawlike generalizations (Adams & Aizawa, 2010). Interestingly, such an argument also undermines Simon's (1981, 1996) idea that there may be sciences of the artificial, including economics and cognitive science. It seems, however, that critics of the extended mind mean the possible lack of unification here: they doubt that there could be cohesive explanations of diverse phenomena. The mere diversity of phenomena, however, or the motley crew nature of artifacts, cannot be in itself sufficient to state that unification is impossible. On the contrary, diversity is the necessary condition to speak of unification in contrast to mere large scope of an explanation. However, unification cannot be ascertained *a priori* – after all, it implies a relation of theory with evidence, as do other virtues in Group 4 in Douglas's helpful taxonomy. It is,

nonetheless, fairly clear what would undermine the motley crew objection: the evidence of invariant structures that support cognitive generalizations.

In summary, both the cognitive bloat objection and the motley crew argument raise concerns about the lack of invariant structures in a theory, rather than simply the overextension of the concept of cognition. Theories that posit complex structures as their basic explanatory entities, but do not remain invariant across different explanations, may be criticized for lacking parsimony. Note that according to Douglas' taxonomy, these criteria for theory choice fall under group 3 and 4, which are considered ideal desiderata for theories. Despite this, explanations, theories, and larger theoretical frameworks can still be scientifically sound even if they lack these criteria. The case of the extended cognition approach appears to fulfill minimal criteria, and there is currently no evidence to suggest a lack of virtues in group 4 (however, the lack of generality may imply that unification cannot be attained).

In general, assessing theoretical virtues in relation to evidence for large theoretical frameworks can be difficult. For example, it can be challenging to assess the scope of a theory without having a definitive list of all cognitive phenomena. A somewhat dated survey lists around 3000 tasks falling into 10 categories, as analyzed in factor-analytic studies, but it has yet to be updated (Carroll, 1993). Performing a systematic survey in reference to all three frameworks of wide cognition is a monumental task that exceeds the scope of this paper.

One point should be stressed, however, before we conclude. It is one thing for a particular extended cognition *explanation* of a specific task to have cognitive virtues, and quite another for the extended cognition *approach* to possess them. As we have called them, extended cognition, distributed cognition, or situated cognition are approaches to the study of cognition. These are not specific and full-blown theories. They could be better thought of in terms of research programs (Lakatos, 1970) or research traditions (Laudan, 1977). Just like other research traditions in cognitive science, such as computationalism or embodied cognition, extended or distributed cognition cannot offer novel predictions for particular phenomena, for example (Wołoszyn & Hohol, 2017). Yet this should not be taken as a sign of their deficiency. Clearly, grand frameworks offer fallible heuristics for building specific theories of particular kinds of phenomena, and these theories may, in turn, inspire fully detailed cognitive models. In the paragraphs above, we have mostly judged the extended cognition view, as opposed to the other research traditions, as a cognitive toolbox that could provide such specific explanations. Quite understandably, these may not exist for a number of phenomena, making the actual scope of cognitive theories in particular approaches fairly limited. For example, most of these frameworks do not seem to be fully worked out for the category of phenomena that Carroll (1993) dubs “cognitive speed”, even if the reliance on external representations and artifacts is usually believed to provide gains in terms of computational efficiency. This, however, implies that there could be significant progress across this category (Vélez et al., 2023).

4. Conclusion

In summary, this paper underscores that the extended cognition approach is not merely a metaphysical perspective on the mind; it comes with its own set of epistemological strengths and weaknesses. While other wide cognition frameworks may exhibit similar internal consistency and empirical adequacy, they diverge concerning the ideal desiderata for theories. Assessing the evidence for each wide cognition approach can be a challenging task, with the extended cognition approach potentially facing greater difficulties in terms of generality and simplicity. This is due to the cognitive bloat objection and the motley crew argument, both of which suggest that there might not be enough invariant structure within the cognitive system to substantiate robust empirical generalizations.

It should be noted, however, that invariance might not be the only way to attain a large breath of scope. Alternatively, the cognitive bloat might not affect the invariance after all, because multiple different cognitive systems could give rise to the same (or sufficiently similar) abstract set of principles that govern them. This is particularly important for the third wave of the extended mind.

In any case, the burden of proof rests with the proponents of the extended framework to demonstrate that their view can meet generality criteria. For instance, a recent contribution by Kirchoff and Kiverstein (2019) links the predictive processing framework, typically regarded as offering grand theoretical unification in cognitive science (Clark, 2016), with the third-wave approach to the extended mind. While the promise of unification remains unfulfilled (Litwin & Miłkowski, 2020), it may not be of particular concern to supporters of the third wave. Unification, after all, is one of the criteria in group 4 for theory choice, representing an ideal desideratum. Currently, the predictive processing approach lacks unequivocal empirical support at its current stage of development (Walsh et al., 2020), making its precise assessment under group 4 criteria unfeasible. From our perspective, relying on another framework with unifying ambitions presents a reasonable strategy for defending the extended cognition view.

In conclusion, our analysis demonstrates that the extended cognition versus situated and distributed views debate extends beyond mere metaphysical or terminological considerations. Paradoxically, this complexity may not bode well for supporters of extended cognition.

Acknowledgements

The authors wish to thank Joel Krueger for comments.

References

Adams, F., & Aizawa, K. (2010). Defending the bounds of cognition. In *The extended mind* (pp. 67–80). MIT Press. https://doi.org/10.7551/mitpress/9780262014038.003.0004

Allen-Hermanson, S. (2013). Superdupersizing the mind: Extended cognition and the persistence of cognitive bloat. *Philosophical Studies*, *164*(3), 791–806. https://doi.org/10.1007/s11098-012-9914-7

Barker, M. J. (2010). From cognition's location to the epistemology of its nature. *Cognitive Systems Research*, *11*(4), 357–366. https://doi.org/10.1016/j.cogsys.2010.05.001

Bernecker, S. (2014). How to Understand the Extended Mind. *Philosophical Issues*, *24*(1), 1–23. https://doi.org/10.1111/phis.12023

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press.

Casper, M.-O. (2023). A Methodological Problem of Choice for 4E Research. In M.-O. Casper & G. F. Artese (Eds.), *Situated Cognition Research: Methodological Foundations* (pp. 17–43). Springer International Publishing. https://doi.org/10.1007/978-3-031-39744-8_2

Clark, A. (2016). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press.

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

Craver, C. F. (2009). Mechanisms and natural kinds. *Philosophical Psychology*, *22*(5), 575–594. https://doi.org/10.1080/09515080903238930

Douglas, H. E. (2013). The Value of Cognitive Values. *Philosophy of Science*, *80*(5), 796–806. https://doi.org/10.1086/673716

Ekman, P., Friesen, W. V., Ellsworth, P., Goldstein, A. P., & Krasner, L. (2013). *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*. Elsevier Science.

Fernandez-Dols, J. M., & Ruiz-Belda, M.-A. (1997). Spontaneous facial behavior during intense emotional episodes: Artistic truth and optical truth. In J. A. Russell & J. M. Fernández-Dols (Eds.), *The Psychology of Facial Expression* (1st ed., pp. 255–274). Cambridge University Press. https://doi.org/10.1017/CBO9780511659911.013

Fridlund, A. J. (1991). Sociality of solitary smiling: Potentiation by an implicit audience. *Journal of Personality and Social Psychology*, *60*(2), 229–240. https://doi.org/10.1037/0022-3514.60.2.229

Keas, M. N. (2018). Systematizing the theoretical virtues. *Synthese*, *195*(6), 2761–2793. https://doi.org/10.1007/s11229-017-1355-6

Kirchhoff, M. D., & Kiverstein, J. D. (2019). *Extended consciousness and predictive processing: A third wave view*. Routledge, Taylor & Francis Group.

Kraut, R. E., & Johnston, R. E. (1979). Social and emotional messages of smiling: An ethological approach. *Journal of Personality and Social Psychology*, *37*(9), 1539–1553. https://doi.org/10.1037/0022-3514.37.9.1539

Krueger, J. (2014). Emotions and the social niche. In C. von Scheve & M. Salmela (Eds.), *Collective Emotions* (pp. 156–172). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199659180.003.0011

Krueger, J., & Szanto, T. (2016). Extended emotions. *Philosophy Compass*, *11*(12), 863–878. https://doi.org/10.1111/phc3.12390

Kuhn, T. S. (1977). *The essential tension: Selected studies in scientific tradition and change*. The Univ. of Chicago Press.

Lakatos, I. (1970). Falsification and the Methodology of Scientific Research Programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the Growth of Knowledge: Proceedings of the International Colloquium in the Philosophy of Science, London, 1965. Vol. 4 Vol. 4* (pp. 91–195). Cambridge University Press.

Laland, K. N., Odling-Smee, J., & Feldman, M. W. (2000). Niche construction, biological evolution, and cultural change. *Behavioral and Brain Sciences*, *23*(1), 131–146. https://doi.org/10.1017/S0140525X00002417

Laudan, L. (1977). *Progress and Its Problem: Towards a Theory of Scientific Growth*. University of California Press.

Litwin, P., & Miłkowski, M. (2020). Unification by Fiat: Arrested Development of Predictive Processing. *Cognitive Science*, *44*(7), e12867. https://doi.org/10.1111/cogs.12867

Menary, R. (2007). *Cognitive integration: Mind and cognition unbounded*. Palgrave Macmillan.

Menary, R. (2015). *Mathematical Cognition: A Case of Enculturation*. Open MIND. Frankfurt am Main: MIND Group. https://doi.org/10.15502/9783958570818

Miłkowski, M. (2018). From Computer Metaphor to Computational Modeling: The Evolution of Computationalism. *Minds and Machines*, *28*(3), 515–541. https://doi.org/10.1007/s11023-018-9468-3

Miłkowski, M., Clowes, R. W., Rucińska, Z., Przegalińska, A., Zawidzki, T., Gies, A., Krueger, J., McGann, M., Afeltowicz, Ł., Wachowski, W. M., Stjernberg, F., Loughlin, V., & Hohol, M. (2018). From Wide Cognition to Mechanisms: A Silent Revolution. *Frontiers in Psychology*, *9*, 2393. https://doi.org/10.3389/fpsyg.2018.02393

Müller, V. C. (2018). The Extended Mind Thesis Is about Demarcation and Use of Words. *Reti, Saperi, Linguaggi*, *2/2018*. https://doi.org/10.12832/92304

Newell, A., & Simon, H. A. (1972). *Human Problem Solving*. Prentice-Hall.

Okasha, S. (2011). Theory Choice and Social Choice: Kuhn versus Arrow. *Mind*, *120*(477), 83–115. https://doi.org/10.1093/mind/fzr010

Rowlands, M. (2009). Extended cognition and the mark of the cognitive. *Philosophical Psychology*, *22*(1), 1–19. https://doi.org/10.1080/09515080802703620

Rupert, R. D. (2009). *Cognitive systems and the extended mind*. Oxford University Press.

Shapiro, L. A. (2011). *Embodied Cognition*. Routledge.

Simon, H. A. (1981). Cognitive science: The newest science of the artificial. *Cognitive Science*, *4*(1), 33–46. https://doi.org/10.1016/S0364-0213(81)80003-1

Simon, H. A. (1996). *The sciences of the artificial*. MIT Press.

Sprevak, M. (2010). Inference to the hypothesis of extended cognition. *Studies in History and Philosophy of Science Part A*, *41*(4), 353–362. https://doi.org/10.1016/j.shpsa.2010.10.010

Sutton, J. (2010). Exograms and Interdisciplinarity: History, the Extended Mind, and the Civilizing Process. In R. Menary (Ed.), *The Extended Mind* (pp. 189–225). MIT Press.

Thagard, P. R. (1978). The Best Explanation: Criteria for Theory Choice. *The Journal of Philosophy*, *75*(2), 76. https://doi.org/10.2307/2025686

Vélez, N., Christian, B., Hardy, M., Thompson, B. D., & Griffiths, T. L. (2023). How do Humans Overcome Individual Computational Limitations by Working Together? *Cognitive Science*, *47*(1), e13232. https://doi.org/10.1111/cogs.13232

Vickers, P. (2013). *Understanding inconsistent science*. Oxford University Press.

Walsh, K. S., McGovern, D. P., Clark, A., & O'Connell, R. G. (2020). Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Annals of the New York Academy of Sciences*, *1464*(1), 242–268. https://doi.org/10.1111/nyas.14321

Wołoszyn, K., & Hohol, M. (2017). Commentary: The poverty of embodied cognition. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00845