Synthese (2023) 202:155
https://doi.org/10.1007/s11229-023-04397-8

ORIGINAL RESEARCH

Rejecting the extended cognition moral narrative: a critique of two normative arguments for extended cognition

Guido Cassinadri[1] · **Marco Fasoli**[2]

Received: 17 March 2023 / Accepted: 17 October 2023 / Published online: 2 November 2023

Abstract

Given the explanatory stalemate between 'embedded' (EMB) and 'extended' (EXT) cognition, various authors have proposed normative and moral arguments in favour of EXT. According to what we call the "extended cognition moral narrative" (EXT-MN) (Cassinadri, 2022), we should embrace EXT and dismiss EMB, because the former leads to morally preferable consequences with respect to the latter. In this article we argue that two arguments following the EXT moral narrative are flawed. In Sect. 2.1 and 2.2, we present respectively King (2016) and Vold's (2018) 'argument from assessment of capacities' (AAC) and Clowes (2013), Farina and Lavazza's (2022a) 'cognitive diminishment argument' (CDA). The AAC states that we should embrace EXT over EMB since the former is better at attributing cognitive credit to individuals with learning disabilities who use assistive tools to complete their learning tasks, thus avoiding their marginalisation. The CDA states that EMB implies a morally undesirable picture of the agent in terms of cognitive diminishment. In Sect. 3, we clarify and criticise the underlying assumptions of the AAC and CDA: the "cognitive credit assumption", the "marginalisation assumption" and, more generally, an ableist conception of disabled agency. In Sect. 3.1, we discuss the role of moto-perceptual skills and metacognitive credit involved in complex cases of tool-use, to demonstrate that the EXT-MN is uninformative in addressing these cases. To conclude, in Sect. 4 we argue that AAC and CDA fail to present EXT as descriptively and normatively superior to EMB.

Keywords Extended cognition · Extended cognition moral narrative · Embedded cognition · Learning disabilities · Cognitive diminishment · Ethics of artifacts

Extended author information available on the last page of the article

1 Introduction

In 1998 Andy Clark and Chalmers (1998) proposed the hypotheses of extended mind and extended cognition (EXT), according to which, under specific coupling conditions between a cognitive agent and a cognitive artifact, the latter can actively and constitutively contribute to the realisation of extended mental and cognitive states, processes, and systems.[1] According to EXT, external scaffoldings in some cases help to produce and explain cognitive phenomena in such a way that they are partly constitutive of the material basis that produces the cognitive processes at stake. Robert Rupert's (2009) embedded view (EMB) admits that external artefactual vehicles can profoundly determine and causally affect the dynamics of neural cognitive processes, by virtue of intense feed-back loops, without constituting extended cognitive processes or systems. According to Rupert, there can be hybrid systems and processes that combine genuinely cognitive internal mechanisms with external non-cognitive resources. EMB admits that environmental factors that causally contribute to the realisation of cognitive processes are explanatorily relevant in order to address certain cognitive phenomena, but this does not force us to be ontologically committed to the cognitive status of the external resources.

Thus, Barker (2010) and Sprevak (2010) argued that although both EMB and EXT posit extracranial mechanisms and processes, the difference is that EMB considers the external elements as merely non-cognitive causal contributors of a hybrid extended mechanism, while EXT considers the external elements as genuine constituents of a higher-level extended cognitive mechanism. For the scope of this article, we use the generic label, 'EMB,' without all the commitments of Rupert's (2009) embedded view of cognition, but rather simply as a view of cognitive integration that does not consider the external artifact as a genuine extended part of the cognitive machinery.[2]

As argued by Barker (2010) and Sprevak (2010), from an empirical, predictive and explanatory perspective it is currently impossible to ascertain when and whether a causal link between the external device and the neural processes (EMB) is a constitutive-extending one (EXT). Given the empirical underdetermination between the two different ontological commitments embraced by EXT and EMB, this dichotomy has resulted in a metaphysical stalemate between the two (Clark, 2011, p. 460Sprevak, 2010; Farina & Lavazza, 2022a, p. 7–8). Moreover, the absence of a universally shared mark of cognition, that defines the necessary and/or sufficient conditions for the external resources to be an extender of a cognitive process (Rowlands, 2010a; Wheeler, 2016, 2019), has further crystallised this stalemate between the two positions.[3]

[1] Given that the difference is not relevant for the purposes of this treatment, we will use the notion of mind and cognition interchangeably.

[2] For example, we will apply Sterleny's (2004; 2010) framework of scaffolded cognition, which considers EXT a limiting case and encompasses most cases of tool-use in EMB.

[3] However, in line with what has been argued by Rowlands, 2010a); Wheeler (2016), 2019), we acknowledge that if it is possible to formulate an encompassing mark of cognition capable of admitting both biological and external components in its definition, then it is possible to postulate extended cognitive systems whenever the external resources respect the mark's necessary and/or sufficient criteria. According to some authors, the predictive processing framework may offer such an encompassing mark of cog-

Given this impasse and the difficulty in preferring either of the two positions from a descriptive and explanatory level of analysis, some supporters of EXT have used normative and moral arguments to support this explanation over EMB. As shown by Cassinadri (2022, p. 3), these authors often appeal to a 'methodological assumption,' according to which the view that reveals itself to be more useful and preferable on an ethical and pragmatic level is also the best description of the state of affairs that we should embrace (Levy, 2007, p. 220; Clowes, 2013, p. 129; King, 2016, p. 45; Vold, 2018, p. 497 Farina and Lavazza, 2022a, p. 8; 2022b, p. 11). We define 'Extended Cognition Moral Narrative', henceforth the EXT-MN, as the argumentative strategy relying on this methodological assumption to support EXT over EMB. The EXT-MN is a general argumentative structure that takes this form:

1. If a view reveals itself to be more useful and preferable on an ethical and pragmatic grounds, then we should embrace it (methodological assumption).
2. Whenever there is a highly integrated cognitive system S, composed of a cognitive agent and a tool, if we embrace EXT, then we obtain moral implication X, while if we embrace EMB we obtain moral implication Y.
3. Since X is morally preferable to Y, we should embrace EXT over EMB to explain S's integration with the tool, even in cases where we have no other well-founded reasons for preferring EXT to EMB.

It is important to note that the EXT-MN works to deliver EXT over EMB only when the system manifests a high degree of cognitive integration (Heersmink, 2015) between the cognitive agent and the artifact. One basic problem of this argumentative strategy is that the different supporters of the EXT-MN do not specify either the conditions of extension or the degree of integration across cognitive dimensions[4] between the agent and the tool (Heersmink, 2012, 2015), which would allow us to consider the system as sufficiently integrated and extensible for moral reasons. Basically, according to the EXT-MN, whenever we are in doubt about whether to interpret a highly integrated cognitive system in terms of EMB or EXT, we should embrace EXT for ethical reasons.

In this article we will discuss some issues arising from this argumentative strategy. We will present and criticise King (2016) and Vold's (2018) 'argument from assessment of capacities' (AAC), in Sect. 1.1, and Clowes' (2013) 'cognitive diminishment argument' (CDA) in Sect. 1.2. These arguments state that, since EMB offers a normatively undesirable picture of the agents' cognitive capacities, we should embrace EXT for normative reasons, by virtue of its capacity to properly attribute cognitive

nition capable of including cases of extended cognitive processes. See Kersten (2022) and Clark (2022) for a recent account of EXT in Predictive Processing. In the event that we reach a universally shared mark of cognition, then the absence or presence of extended cognitive systems and processes would become an empirical matter. Thus, this solution might be sufficient for settling the debate between EMB and EXT in controversial cases of highly integrated cognitive systems and EXT-MN would be superfluous.

[4] Heersmink (2012, 2015) offers a rich, nevertheless not complete set of dimensions of cognitive integration between the agent and the tool, considering the dimensions of reliability, durability, trust, procedural and representational transparency, individualisation, bandwidth, speed of information flow, distribution of computation, and cognitive and artifactual transformation.

credit and agency to individuals. In Sect. 2.1, we present and criticise the assumptions on which the AAC and CDA are based, which are the "cognitive credit assumption", the "marginalisation assumption" and the commitment to an ableist conception of cognitive disability and disabled agency. In Sect. 2.2 we discuss the role of moto-perceptual skills and metacognitive credit in tool use, which are ignored by the "cognitive credit assumption" in the attribution of credit to individuals in learning settings. To conclude, in Sect. 3.1 we argue that the AAC fails to present EXT as a view that offers a less marginalising picture of cognitive disability.

In summary, although our article must be intended as a critique of two strong arguments of the EXT-MN, namely AAC and CDA, rather than a critique of EXT itself, our analysis of these arguments reveals some weaknesses that prompt us to embrace EMB. Thus, our arguments begin by assuming the presence of a metaphysical and explanatory stalemate between EMB and EXT for explaining highly integrated cognitive systems, as assumed by some supporters of the EXT-MN (King 2016, p. 44); Farina & Lavazza, 2022a, pp. 7–8). We then criticise the use of two normative arguments in favour of EXT. To conclude, in the article, we also aim to show that the range of possible relationships between external artifacts and the human mind is complex and heterogenous, and cannot be addressed in the way it is done by the AAC and CDA.

2 The argument from assessment of capacities and the cognitive diminishment argument

2.1 The argument from assessment of capacities (AAC)

Before presenting the 'argument from assessment of capacities' (AAC) as it is formulated by its two major proponents, Caroline King (2016) and Vold (2018), it is important to clarify the theoretical differences between the overall frameworks endorsed by the two authors. Concerning the metaphysical difference between EMB and EXT, King (2016, p. 43) appeals to Ladyman and Ross' (2010) stance, according to which the metaphysical debate regarding the location and spatial extension of cognitive processes is misguided and affected by a flawed metaphysics of 'containment', typical of an immature science. Thus, she remains metaphysically agnostic and consequently considers the explanatory arguments in favour of either of the two theories to be irrelevant. Her argument strategy follows the 'methodological assumption', by embracing EXT over EMB for its normative and moral benefits. She first considers some normative claims concerning cognitive disability and cognitive credit, arguing that "if they are claims that we find persuasive, and if the extended mind best captures them, then we have good normative reasons to adopt the extended mind framework" (King, 2016, p. 45).

Vold (2018) instead combined both explanatory and normative reasons to support EXT over EMB, refusing to assume the presence of an explanatory indistinguishability between the two. First, she argues that EXT is explanatorily better than EMB by virtue of its (1) simplicity, (2) usefulness, and (3) stronger explanatory power (Vold, 2018, pp. 493–496). However, since in this article we prefer to maintain a narrow

scope, we will not focus on her explanatory arguments,[5] and we will criticise her use of the normative 'argument from assessment of capacities' (AAC). This argument, proposed by both King (2016) and Vold (2018) (henceforth K&V when they present overlapping arguments), states that EXT is morally preferable to EMB because the former better assesses the capacities of learning-disabled (LD henceforth for learning-disabled) individuals who rely on compensatory tools to complete their learning tasks, thus avoiding their marginalisation.

AAC: given that EXT better assesses the capacities of LD individuals who rely on compensatory tools to complete their learning tasks than EMB, offering a less marginalising picture of cognitive disabled individuals, EXT is morally preferable to EMB, and we should embrace EXT.

First, K&V implicitly assume what we call the "cognitive credit assumption", according to which what matters on normative grounds in educational settings is reduced to the degree of cognitive credit performed by the agents. The notion of 'cognitive credit', introduced by Vold (2018, p. 501), refers to the degree (more or less) of cognitive capabilities performed by the agents (King, 2016, p. 56; Vold, 2018, p. 501).

According to King (2016, p. 44) and Vold (2018, p. 502), the correct attribution of cognitive credit has direct moral implications, in the form of normative claims concerning the relation between cognitive credit and marginalisation. King's commitment to the EXT-MN is made explicit by her claim that the way in which we conceptualise cognitive diversity has relevant ethical consequences, especially concerning the marginalisation of cognitively atypical individuals (King 2016, p 44). In fact, according to King, by attributing a lesser degree of cognitive credit to an individual with cognitive disability, we are contributing to her morally undesirable depiction of being capable of less, and to her morally undesirable condition of being a marginalised agent within society. We call this the "marginalisation assumption". Thus, King argues:

my project is not to defend the truth or desirability of these normative claims; I will simply explore whether the extended or embedded cognition framework best captures them. Then, I will argue that if they are claims that we find persuasive, and if the extended mind best captures them, then we have good normative reason to adopt the extended mind framework. (King 2016, p. 44)

[5] Nevertheless, we think that her arguments fail to properly respond to Sprevak (2010) and Barker's (2010) analysis, which highlight that from an explanatorily, empirical, predictive, and experimental perspective it is almost impossible to distinguish between EMB and EXT. Barker (2010, pp. 364–365) shows that the appeals in the literature to 'simplicity' in favour of EXT are usually brief, vague, and ultimately inconclusive. Moreover, it is not clear whether EMB or EXT is the simpler explanation. In fact, EMB adds more psychological predicates ('transcranial kinds'), sparing the ontological commitment, while EXT offers a simpler explanation at the cost of a heavier ontological commitment (Sprevak, 2010, p.17). Vold appeals to Levy's considerations (2007, p. 220), arguing that EXT is useful for explaining how external scaffoldings can support internal willpower in cases of addiction. However, neither Vold nor Levy offer clear criteria for distinguishing environmental scaffolding from mind-extenders, and this may lead to a cognitive bloat (Adams & Aizawa, 2008); namely, the excessive proliferation of the mind into the environment (Cassinadri, 2022, p. 12). To conclude, Sprevak (2010) demonstrates that the explanatory power of EMB and EXT is equivalent, as in both cases the explanatory work is done by 'extracranial' kinds to explain hybrid cognitive operations, independently of the mental/non-mental ontological commitment on the external resource.

Therefore, the AAC represents a specific instantiation of the EXT-MN, by arguing that we should embrace EXT, since EMB fails to attribute the correct dose of cognitive credit to LD individuals who rely on compensatory tools to complete their cognitive tasks, thus offering a marginalising picture of their condition. King is committed to the methodological assumption, since she states that "if the extended cognition framework offers a way of conceptualising cognitive diversity that is less marginalising of LD individuals than an intracranial framework, we have serious ethical reasons to consider adopting it" (King, 2016, p. 44).

To argue this, K&V first distinguish two kinds of remedial strategies for cognitive rehabilitation, which are 'restorative' and 'compensatory' (King 2016, p. 53; Vold, 2018, p. 498). The former aims at restoring the damaged neural area, whereas the latter aims to achieve the same functional results as restorative strategies by combining neural processes with external resources. They then draw an analogy, by arguing that the same distinction applies to the learning strategies used for addressing learning difficulties (Garner & Campbell, 1987), namely the 'remedial' and 'compensatory' approaches (King, 2016, p. 54; Vold, 2018, p. 500). The former attempts to directly address the individual's impairment and improve her ability to perform a particular task in the same way as a non-disabled individual, whereas the latter attempts to 'circumvent' the impairment and help the individual perform a task using assistive technology (King, 2016, p. 54). At this point, the AAC takes this form:

If we are to understand remedial strategies for LD as being analogous to restorative strategies in cognitive rehabilitation, the embedded mind thesis suggests that remedial strategies are the only way to truly *enhance* an LD individual's cognitive capabilities. The compensatory approach, on the other hand, is simply a way of *helping* the individual get around her impairment and complete the task. However, the extended mind thesis allows us to say that even compensatory strategies for addressing learning disabilities are genuine ways of *improving* or *increasing* an LD individual's cognitive capabilities, not simply circumventing an impairment. (King, 2016, p. 54, our italics)

Thus, following this line of reasoning, if we embrace EMB to assess the capacities of these individuals, then we diminish both their cognitive agency and their cognitive credit by reducing these cognitive agents to be considered as mere tool-users. Instead,

if the bio-external scaffolding that LD individuals who use compensatory assistive technologies use can, in some cases, be considered to be part of their cognitive systems, then an individual using such technologies has, in a robust sense, improved cognitive capabilities. (King, 2016, p. 54)

Thus, K&V's conclusion is that EMB fails to properly attribute cognitive agency and 'cognitive credit' (Vold, 2018, p. 501) to individuals with LD who are supported by an assistive tool, whereas EXT properly acknowledges the cognitive agency and cognitive credit of these individuals, by considering the functional work performed by the latter as constitutive parts of the individuals' extended cognitive system. Thus, it is important to stress that K&V's use of the notion of cognitive credit assumes both a descriptive and a normative meaning: the first refers to the amount of cognitive work performed by the agent, while the latter refers to the blame and praise that the agent deserves according to the amount of cognitive work she is responsible for.

Let us first clarify the role of the descriptive notion of cognitive 'credit' within the AAC. K&V define EMB in terms of the 'principle of intracranialism' (PI), according to which there is an.

inverse relationship between the extent to which an individual relies on external scaffolding and the extent to which we ought to say that her mind is really doing x, where x is some cognitive process. […] neural activity thoroughly determines mental activity. (King, 2016, p. 55; Vold, 2018, pp. 17–18)

K&V argue that since EMB is committed to PI, EMB fails to deliver the right amount of normative credit to those LD agents who use compensatory tools in learning settings. Thus, AAC uses a normative concern to impose a specific description of the states of the world (EXT) over another (EMB), using the argumentative pattern of the EXT-MN. According to AAC, given that EMB assigns a limited amount of cognitive work to the individual by reducing it to her neural resources (PI), an evaluation of her cognitive activity using EMB fails to respect the normative value of 'fair attribution of cognitive credit' by judging her as doing less rather than equal work of her peers who complete the same task by virtue of their neural resources only.

If that's the case, then an LD individual who relies on assistive technology to complete cognitive tasks is only "doing" as much as her neurons are doing. That is, if her cognition is exclusively located in her neurons, then within a given task, whatever work is being done by the bio-external tools is not done by her. The more heavily integrated the assistive technology into a cognitive process, the less cognition the person is performing. Thus, if the embedded thesis is right, it commits us to saying that LD individuals who rely heavily on assistive technologies are cognitively capable of less than non-LD individuals. But this seems wrong. Assistive technologies are tools that help LD individuals do *more*, not less. (King, 2016, p. 56, our italics)

King draws a dichotomic scenario: either we say that the tool is part of the agent's cognition, and therefore we can say that LD individuals who use compensatory tools are cognitively capable of more, or we are offering a marginalising picture of them as mere tool-users and cognitively capable of *less*. Thus, by assuming that we have an ethical reason for adopting the less marginalising framework (King 2016, p.44), and by assuming that the attribution of a less degree of cognitive credit to an individual implies a morally undesirable picture and condition (marginalisation) of such an individual, she concludes that we should embrace EXT for ethical reasons.

King discusses the case of Dana as an application of the AAC. Dana is a student who uses a graphic organiser to aid her strategic planning and decision making, enabling her to complete cognitive tasks that she could not otherwise perform. According to King, it would be unfair to say that Dana is unable to perform these cognitive operations just because she needs an external support tool to complete them. Since EMB conceptualises the graphic organiser as an external tool which is not part of Dana's cognitive system, by embracing EMB we consider the organism-bound agent, Dana, as incapable of performing specific tasks. Thus, since EMB fails to correctly attribute cognitive credit to Dana, we have a good normative reason to embrace EXT to explain such a case. To conclude her argument, King also appeals to the reported feelings of well-assisted LD individuals, who feel more confident, more independent, and more capable when using assistive technologies (Young and Specht

2010; 2011), as a proof of the validity of EXT. She uses these empirical, subjective, and phenomenological findings, arguing that.

if the extended mind hypothesis is correct, and therefore cognitive capabilities are attributed to extended organism-plus-environment systems, then it is true that the assistive technologies help the individual do more rather than less. (King, 2016, p. 59)

Thus, she demonstrates in this way that EXT offers a less marginalising and more inclusive view of cognitive diversity, when compared with EMB. Vold (2018) simply appeals to King's argument, concluding, as she does, that by embracing EMB we would consider Dana as incapable of making complex decisions, and thus this framework fails to accommodate the supposedly correct intuitions about LD individuals (Vold, 2018, p. 500).

To sum up, in this section we introduced the AAC, which has been defended in slightly different forms by the supporters of the EXT-MN King and Vold. We argued that AAC implicitly assumes both what we call the cognitive credit assumption, according to which what matters on normative grounds in educational settings is reduced only to the degree of cognitive credit, and the marginalisation assumption, that states that attributing a low degree of cognitive credit to an agent means marginalise her/him. Moreover, we showed that according to K&V EMB is committed to PI, the principle that - to put it loosely - states there is an inverse relationship between the cognitive credit of an agent and the employment of external artifacts.

2.2 The cognitive diminishment argument

Before discussing the AAC argument and its problems in details, let us consider the cognitive diminishment argument (CDA). CDA can be considered as a generalisation of K&V's AAC, since its scope is not reduced to the case of individuals with cognitive deficit or learning disabilities, but rather applies to any scenario in which the reliance on specific tools may lead to the enhancement or diminishment of specific cognitive capacities.[6]

The fact that the tools we use may eventually lead to the cognitive diminishment of the tool- user is a problem that has its first philosophical formulation in Plato's *Phaedrus*, where Socrates presents the potentially undesirable implications of relying on writing and reading rather than on biological memory. He argued that the written language would undermine and diminish biological memory as it allowed us to store information in the environment, rather than in the brain, subsequently making us cognitively lazy (Heersmink, 2017, p. 27). However, the introduction of writing systems allowed humanity to rely increasingly upon epistemic artifacts to complete complex cognitive operations that would be impossible without the use of external scaffoldings, thus enabling the construction of a complex civilization (Wolf & Stoodley, 2008). The more the computational power of our portable devices increases, the more we tend to delegate complex operations to them, and this may lead to inter-

[6] Moreover, similarly to the AAC, which was originally formulated by King (2016) and then simply reproposed by Vold (2018), the CDA that has been formulated by Clowes (2013) and then Farina and Lavazza (2022a) simply appealed to his argument.

related phenomena of cognitive enhancement and cognitive diminishment that are complex to morally assess and evaluate (Heersmink, 2017, p. 27).

For example, maps are a kind of cognitive artifact that have traditionally been used to enhance cognitive orientation. The proper functioning of these tools requires the agent to perform complex orientation and perceptual-coordination operations, and for this reason only a few people are still using maps instead of GPS devices to orient themselves. Although the functional use of GPS still requires us to deploy our cognitive, perceptual, and motor capacities, they lighten our cognitive load when compared with maps (Fasoli, 2016, pp. 69–70). However, the human sense of orientation is a cognitive capability that needs to be practised in order to be kept in good condition (Woollett & Maguire, 2011). Thus these devices, along with many others, raise the same issue stressed by Socrates in the *Phaedrus*: are we going to lose specific capacities by intensively delegating complex operations to external tools? The issue has been raised concerning the potential risks of intensive use of the internet. For example, Carr (2008, 2010) argues that the pervasive reliance on internet functionalities is making us more distracted, lazy, and less capable of critical thinking. According to Clowes (2013), this picture of humanity, as increasingly diminished and impaired by the tools we use, for instance by the so-called Google Effect (Sparrow et al. 2011), is a direct consequence of an internalist picture of the mind.

From the internalist (and embedded) vantage-point it is as if our minds are being steadily off-loaded and dissipated into our tools, potentially with the result that we become *less autonomous*, less knowledgeable and arguably *less interesting* sorts of creatures in the process […] For EMB (or Brainbound) theorists, it follows that if our minds are being bled out into a series of environmental resources we are *diminished* in the process; (or so it seems at first pass). For EXT theorists this does not follow, for rather than being *dissipated* into our tools, we are incorporating them into us.[7] (Clowes, 2013, pp. 127–128, our italics)

Then, in line with the EXT-MN, Clowes argues that, given the ethical implications of EXT and EMB, we have a good reason to prefer the former over the latter (Clowes, 2013, p. 129).[8]

The framework he is using assumes that every time we delegate cognitive operations to external devices, we are losing a part of ourselves. EMB is associated with

[7] Clowes originally used the labels 'HEMC' for the embedded cognition and 'HEC' for extended cognition.

[8] "It starts to become apparent that these positions—perhaps unexpectedly—present *ethical implications* (in the wide sense) for how we valorise cognitive changes that may take place and have implications for how we consider human cognitive futures. This brings us to *one reason to prefer* EXT over EMB. EMB seems to offer us a rather limited viewpoint on the incorporation of E-Memory which points toward (if not necessitating) cognitive diminishment; EXT seeming to imply the deep incorporation of at least E-Memory technologies involves knowledge being bled out of agents. EXT theorists on the other hand have a more nuanced set of positions they can use to qualitatively assess, e.g. what boundaries make an agent look most coherent? As The Cognitive Integration of E-Memory, the EXT approach could make available an expanded set of theoretical tools to examine the new unities which might arise around E-Memory it appears to be a *more useful lens*" (Clowes, 2013, p. 129 our italics). In this quote Clowes clearly appeals to the criterion of usefulness and to an ethical reason to prefer EXT, given its ethical implications. Thus, although Clowes is not explicitly committed to the EXT-MN and to the methodological assumption as we have presented, he is committed to an implicit and informal version of both.

de-skilling, diminishment, and dissipation of *agents* and of their capabilities. Thus, cognitive delegation and offloading is not characterised in relational terms as a dynamic process in which the agent is actively engaged with an external scaffolding, but rather is presented simply as a passive operation in which the agents are deprived of crucial parts of themselves. As King (2016) and Vold (2018) would put it, cognitive delegation and cognitive offloading, framed in EMB terms, consider agents as doing and being 'less', rather than 'more'.

The identification of what is enhanced and what is diminished, namely what is increased and what is reduced, depends on the boundaries of the agent. Thus, according to Clowes,

Brainbound theorists get the limits of the agent wrong and in so doing can miss interesting aspects of the total agent […] The Extended theorist is interested in the wider ensemble of human + that is, the agent plus the motley collection of equipment which has been densely incorporated (and some of that which is only shallowly incorporated too). (Clowes, 2013, p. 128)

Thus, even the cognitive diminishment argument (CDA) assumes a dichotomic way of thinking, according to which, either we frame cognitive delegation in terms of the extension of the agent and the consequent enhancement of her capacities, or we picture the agent in intracranialist terms as cognitively diminished every time she delegates parts of her cognitive operations to external supports. This dichotomic framework is also used by Farina and Lavazza, who are explicitly committed to the EXT-MN (Farina & Lavazza, 2022a, p. 8; 2022b; 2022c, p. 11). In fact, before presenting their version of the CDA, they acknowledged that there is a descriptive and metaphysical stalemate between these competing views and that this guided their choice to use normative and moral arguments to overcome the impasse.

However, if there is no way — in metaphysical terms — to adjudicate the debate in favour of one of the two theses we discussed, we can *assume* that our moral preferences might guide our choice. In this case, we can figure out what the moral consequences of adopting each view would be. However, what appears ethically preferable to us can be also taken as the description of the states of fact in the world, although we do not have the epistemic certainty that they correspond to our moral preferences. (Farina & Lavazza, 2022a, p. 8).

Once they made their commitment to both the methodological assumption and the EXT-MN clear, they appealed to Clowes's argument to support EXT over EMB. According to them, by embracing EXT,

we are not being *dissipated* into our tools and therefore made less efficient; instead, we *actively* incorporate the tools we use into our cognitive repertoire, and in doing so, we can become *smarter*. (Farina & Lavazza, 2022a, p. 11 our italics)

Note again that the vocabulary highlighted in italics reveals the dichotomic distinction between active and passive, empowered and diminished. To conclude, according to the framework assumed by the CDA, by using a tool the agent is passively dissipated, while the enhancement of agency by virtue of cognitive extension is an active operation. Since the CDA is framed as a moral argument, the fact that EMB implies a diminishment of the agent is not expressed merely in descriptive terms, but rather in normative ones, by considering such a diminishment as a morally undesirable picture and a characterisation of human agents as less efficient, passive and stupid.

However, we think that the CDA leads to this conclusion only by assuming a specific moral framework, which is crucial for the argument. We think that the moral desirability or undesirability of the implications of the two views (EMB/EXT) cannot be derived merely and directly from the causal/constitutive cognitive distinction, but rather from the combination between this distinction and a specific moral framework which is implicitly assumed. This in line with Cassinadri's (2022) general critique of the EXT-MN, according to which, if we consider the specific arguments in favour of the EXT moral narrative proposed so far, it becomes clear that the supposed relevant moral differences and implications attributed to the two views are not the result of the two views taken per se, but rather are the result of further moral assumptions and frameworks that are used in combination with EMB and EXT.[9] Therefore, in Sect. 2.1, we will argue that both the AAC and the CDA assume a specific moral assumption on which the normative considerations in favour of EXT ultimately depend.

3 The theoretical assumptions underlying the AAC and the CDA

The AAC is based on two fundamental assumptions that we mentioned in Sect. 1.1 above, namely the "cognitive credit assumption" and the "marginalisation assumption". The former considers only the degree of cognitive credit of the agent as normatively relevant in educational settings.[10] According to the latter, a lesser degree of cognitive credit implies a morally undesirable picture of the agent's worth: the less cognitive credit is attributed to a cognitive agent for performing a task by using a tool, the more this implies her marginalisation. First of all, it seems that a major problem in K&V's argument is the insufficient or lack of justification for both the "cognitive credit assumption" and for the "marginalisation assumption", which are simply implicitly assumed but not discussed by the authors.[11]

Second, it becomes clear that the AAC presents EXT as morally preferable to EMB only by virtue of these two specific assumptions. Thus, the real structure of the AAC takes this form of the EXT-MN:

1) If a view reveals itself to be more useful and preferable on an ethical and pragmatic grounds, then we should embrace it (methodological assumption),
2) What is normatively relevant in educational settings is reduced to the cognitive credit of agents (cognitive credit assumption).

[9] Cassinadri (2022) criticised the argument from better protection, the argument from legitimate interventions, the argument from extended agency and autonomy and the argument from better treatment, arguing that each of them present EXT as morally preferable than EMB not for reasons inherent to the two frameworks taken per se, but rather by virtue of specific moral assumptions.

[10] As we will see in the next section, this assumption ignores the normative relevance in educational settings of moto-perceptual skills (Sterelny, 2004, 2010) and of metacognitive credit (Heersmink and Knight 2018).

[11] King (2016, p.44) simply mentions the fact that the clinical study of cognitive diversity has led to the marginalisation of atypical individuals. However, this does not constitute a sufficient justification of the marginalisation assumption.

3) The less cognitive credit can be attributed to an agent, the more she/he is marginalised (marginalisation hypothesis).
4) If we embrace EXT, then extended cognitive agents are not marginalised, because we can attribute to them the same amount of cognitive credit we attribute to neurotypical agents.
5) If we embrace EMB, then embedded cognitive agents (cognitive tool-users) are marginalised, because in so doing we attribute less cognitive credit to them than to neurotypical agents.
6) Thus, we have a normative and ethical reason to prefer EXT over EMB for explaining how agents with learning disabilities use compensatory tools for performing learning tasks.

As already anticipated in the last section, in line with what has been argued by Cassinadri (2022), it is clear that the EXT-MN presents EXT as morally preferable then EMB only by virtue of specific moral assumptions, according to which we can extrapolate the different moral implications of EXT and EMB. Thus, the moral preferability of the consequences of the two different views EMB/EXT depends primarily on the moral framework that is assumed. In this case the role is played by the "marginalisation assumption". Thus, in order to use the EXT-MN as a general argumentative strategy to overcome the EXT/EMB stalemate in favour of EXT, the supporters of this narrative should first provide independent reasons for accepting the moral framework according to which EXT implies morally preferable consequences compared with EMB. The AAC instead works by first assuming a marginalising picture of disability (marginalisation assumption), K&V then use EXT to overcome the marginalisation by extending the boundaries of the agential system.

Now that we have clarified this point, we will focus on the common moral assumptions underlying both the AAC and the CDA. They both assume that an agent who is capable of completing a task solely by virtue of her cognitive capacities should be considered to be in a more morally desirable condition compared with an agent who completes the same task by relying on an external tool that is not part of the agent's cognitive system, thus deserving less cognitive credit. The first kind of agent is considered by King, Vold, Clowes and Farina and Lavazza as independent and active, while the dependent agent is considered as passive and dependent. The AAC and CDA attribute a moral value to the difference between active and passive agents, by considering the former picture of agency as more valuable and morally desirable compared with the latter. This common assumption seems to be based on a specific conception of human agency that morally evaluates independence compared with dependency. This account of agency falls within the notion of 'agential bias'; namely, the idea that the distinctive feature of human dignity is defined by the capacity of acting and causing things, rather than being caused and being acted upon (Reader, 2007). Agential bias evaluates persons as beings capable of actively and independently choosing to do what they do by virtue of exercising their intrinsic capacities. 'Independence' captures the idea that the agent is not caused to act by external ele-

ments, making the agents' rational decision sufficient to actualise their capabilities (Reader, 2007, p. 587).[12]

More generally speaking, we believe that the definition of agency employed by King (2016), Vold (2018); Clowes (2013); Farina and Lavazza (2022a) assumes an ableist conception of human agency, by considering the relational and dependent character of disabled human agency as something as intrinsically morally undesirable. Ableism can be characterised as a set of beliefs and practices that considers and produces disability as a counter-image to able-bodiedness and thus as deviance, unwanted difference and an unworthy condition (Tarvainen, 2019; Campbell, 2009: 3–30, 197; Ahlvik-Harju, 2015).[13]

The AAC and CDA first assume an ableist picture of human agency and then propose to overcome the supposed inherent unworthiness and marginalisation of the disabled condition by extending the boundaries of the agential system through cognitive extension, which is framed in terms of an enhancement of the individual's agency and capacities while they assume a morally undesirable picture of an agent who completes a cognitive task by relying on external support. In this case, an ableist conception of agency contains the inherent moral assumption that relational and dependent agents are necessarily more marginalised (AAC) and diminished (CDA), thus motivating King (2016), Vold (2018), Clowes (2013), and Farina and Lavazza's (2022a; Farina & Lavazza, 2022b, c) moral urge to extend the boundaries of such agents by expanding their cognitive system. So the AAC and CDA follow this typical pattern of the EXT moral narrative:

1) If a view reveals itself to be more useful and preferable on ethical and pragmatic grounds, then we should embrace it (methodological assumption).
2) Disabled agency (AAC) and dependent agency (CDA) are forms of impaired agency that are morally undesirable and less morally worthy (ableist conception of human agency).
3) Extended cognitive agents postulated by EXT are active and independent because they complete cognitive tasks without relying on external support.
4) Cognitive tool-users depicted by EMB *passively* rely and depend on external support to complete their cognitive tasks.
5) EXT implies a conception of cognitive agency and cognitive disability through tool-use which is morally preferable compared with the one that is implied by EMB.
6) We should embrace EXT over EMB.

In this case, it is the ableist conception of human agency that operates as the ad-hoc implicit moral assumption that allows these authors to present EXT as morally

[12] Similar to Reader, Francis and Silver (2007), who argue that independence is the crucial moral feature that liberalism attributes to individuals who are capable of rationally self-determining their own conception of good.

[13] "Ableism is the set of ideas, practices, institutions and social relations that presume ablebodieness, and by doing so, construct persons with disabilities as marginalized… and largely invisible others" (Chouinard, 1997, p. 380).

preferable to EMB.[14] Thus, the AAC and the CDA generally work by first assuming an ableist framework that morally devalues disability, dependency, and relational agency. They then use EXT to overcome the marginalisation and diminishment inherent in a condition of dependency on an external support by extending the boundaries of the agential system through the extension of cognition.

Although it is not possible here to fully justify an alternative relational, dependent, and ecological account of human agency, which does not devalue the agent's dependence on external support (Mackenzie & Stoljar, 2000; Francis and Silver, 2007; Timpe, 2019), we simply suggest that once we abandon the fundamental assumption on which both AAC and CDA are based, these arguments no longer demonstrate that EXT is morally preferable to EMB. We think that there are good reasons for abandoning this framework of agency, because it fails to acknowledge that agency and patiency are correlates, rather than opposite aspects of human action (Reader, 2007) and cognition, and that this relational, dependent, and supported cognitive agency is not necessarily less valuable and worthy. Thus, in line with Reader (2007), as well as with Francis and Silver (2007), we consider the ableist conception of agency assumed by the AAC and CDA as flawed, because it wrongly ignores and obscures that the independence of paradigmatic rational human agents depends on a history of support and scaffolding from their socio-material, affective, and cognitive (Sterelny, 2010) environment. Thus, this conception wrongly undervalues the variety of ways in which human agency is ecologically supported and dependent on external resources (Timpe, 2019), as well as the complex set of skills and capabilities via which the agent can interactively rely on, act, and react to his external influences and resources, whether these are social (Mackenzie & Stoljar, 2000), or technological (Fasoli, 2018).

Although this point has also been stressed within disability studies (Brownlee & Cureton, 2009), this conceptualisation of human agency inclusively accounts for both disabled and non-disabled individuals. Relational, ecological, and dependent accounts of agency stress that agency should not be reduced merely to the individual capacity of internally generating intentional actions and reasonable choices, since it is often socially and technically scaffolded, mediated, and transformed (Mackenzie & Stoljar, 2000; Silver and Francis, 2007; Timpe, 2019).

3.1 Cognitive credit, moto-perceptual skills and meta-cognitive credit

In this section we will consider what is left out by the "cognitive credit assumption" when we normatively evaluate the credits of individuals in learning settings, namely the role of moto-perceptual skills and the metacognitive credit. Before defining these notions, let us first return to the 'principle of intracranialism' (PI), as formulated by King (2016) and Vold (2018). This principle is attributed to EMB and it establishes

[14] To present a similar case, Cassinadri (2022, pp. 13–14) also criticises Farina and Lavazza's (2022a, p. 14) 'argument from better protection', which relies on the same argumentative strategy typical of the EXT narrative. They first assume that full moral status is granted to individuals with sophisticated cognitive capacities and then argue that EXT guarantees a better protection to cognitively dependent agents. However, EXT offers better protection to these individuals only by assuming that non-extended cognitively deficient agents are not granted full moral status, leaving open the so-called 'problem of marginal cases' (Wasserman et al., 2017).

the presence of an inverse relationship between the extent to which an individual relies on external scaffolding and the extent to which we ought to say that her mind is really doing x, where x is a cognitive process. (King, 2016, p. 55; Vold, 2018, pp. 17–18). We acknowledge that EMB is often committed to PI and therefore implies a lesser degree of cognitive credit attributable to a tool-user when compared with EXT. Nevertheless, we think that the real problematic hypothesis is not PI but the "cognitive credit assumption" embraced by K&V, which states that what matters on normative grounds in educational settings is reduced to the degree of cognitive credit performed by the agents. This fails to capture the entire picture of what is normatively relevant in skilled cognitive tool-use, namely moto-perceptual skills (Sterelny, 2004, 2010) and the degree of meta-cognitive credit (Heersmink and Knight 2018), by focusing the difference between EXT and EMB merely in terms of the degree of use of cognitive capacities. The notion of metacognitive credit captures the degree and proficiency of exercise of the metacognitive capacities performed by the agent. Metacognition, classically thought of as "knowing about knowing" (Metcalfe & Shimamura, 1996) is the capacity to think about one's own cognition. Here by metacognition we mean the agent's awareness of the availability of both internal and external resources, and of capacities for efficiently using such resources for the completion of the relevant cognitive task at hand (Heersmink and Knight 2018). Developing and performing this kind of metacognitive awareness is a relevant learning goal (Kuhn, 2000) which is ignored by K&V's "cognitive credit assumption".

Moreover, moto-perceptual skills are those capacities that are involved in skilled tool-use of epistemic and cognitive artifacts, through which "we transform difficult cognitive problems into easier perceptual problems" (Sterelny, 2004, p. 7). Dana's graphic organiser is a perfect case of transformation of a purely cognitive problem (decision making), into a cognitive, motor and perceptual one (skilled use of the graphic organiser). By using EMB we can acknowledge the presence of specific skills and practices through which Dana uses her graphic organiser to undertake complex processes of decision-making. Thus, despite the fact that EMB attributes less cognitive work and cognitive credit to Dana, this does not have absolute value insofar as if we abandon the cognitive credit assumption, we can consider the moto-perceptual skills performed by the agent as normatively relevant in educational settings. We think that we have a good reason for abandoning such an assumption, because although relying on external support for 'completing a cognitive task x' may imply a decrease of cognitive credit, in the EMB perspective this reliance is not merely a form of 'not doing x', i.e., a passive delegation of that part of the task, as suggested by K&V. Rather, some forms of cognitive delegation should be understood as forms of active moto-perceptual and cognitive practices, which require a series of complex cognitive and sensorimotor skills through which the agent offloads specific functions to the artifact, which is integrated to his internally reconfigured cognitive capacities. As argued by several authors with different theoretical commitments (Fasoli, 2018; Heersmink, 2012; Hutchins, 1995; Menary, 2018; Sterelny, 2004, 2010), the skilful and expert use of cognitive artifacts depends on and fosters the emergence of a new set of sensorimotor and cognitive capacities, through which internal cognitive processes are reconfigured, forming new hybrid forms of cognitive integrated systems (Fabry, 2018). Moreover, "cognitive artifacts often not only interact with

one of our cognitive abilities at a time but may engage our cognitive system in many ways simultaneously" (Fasoli, 2018, p. 679), in such a way that "the acquisition of cognitive practices leads to the transformation of our cognitive abilities" (Menary, 2007). The transformative integration of the cognitive and moto-perceptual capacities of the agent and the tool can be framed and explained both in terms of extended (Menary, 2007) or scaffolded (Sterelny, 2004, 2010) cognition, which can be traced back more generally to EMB. The main difference is that in the latter case the emergence of new capacities within an integrated cognitive system does not imply that the external resource becomes a constitutive component of the cognitive processes of such a system. Nevertheless, as long as we do not reduce what is normatively valuable in educational settings only to the degree of cognitive credit attributable to the system (cognitive credit assumption), we can attribute and ascribe the right degree of normative credit to individuals, who are not extended cognitive systems that perform skilled moto-perceptual use of tools and exercise their metacognitive capacities.

To offer other examples of cognitive tool-use, consider Dana's colleague Ava, who suffers from the same LD. She doesn't use a graphic organiser but an A.I. software such as ChatGPT, to which she constantly delegates her strategic planning and decision-making. Every time she wants to make a decision or to plan something, she consults the software for guidance, completely delegating her decisions to it. Now, neither King (2016) nor Vold (2018) discuss the criteria according to which an artifact is "deeply integrated" with a user's mind, thus if we consider the dimensions identified by Heersmink (Heersmink, 2015, Heersmink and Knight, 2018), namely the information flow, accessibility of the artifact, trust, procedural transparency, informational transparency, distribution of computation, and personalisation, it is plausible that under specific circumstances, despite not defined by K&V, according to K&V the software could be sufficiently integrated to the agent so as to allow us to posit an extended cognitive system for normative reasons, as in the case of the graphic organiser. The AAC would consider the case of Dana and Ava as two instances of extended cognitive agents, who perform the same degree of cognitive credit for completing the same cognitive task. However, in these two cases the moto-perceptual skills involved, the degree of brain-based cognitive credit and the degree of metacognitive credit attributable to the two agents could significantly differ. Should we really attribute to Ava the same cognitive credit as her colleague (see also Marconi, 2005)? Given that the outputs of the extended systems are the same, should we really consider the differences between the two systems as irrelevant? Both these possibilities would have several counterintuitive consequences.

For instance, in education it would be impossible to attribute higher cognitive credit to students who use their brains plus pen and paper to do a column multiplication than to those that use a digital device for the same task. For the same reason, it would also be impossible to consider the latter as cheating, and this appears quite controversial. Even more controversial is the fact that, according to the EXT moral narrative, the A.I. software improves Ava's cognitive abilities 'in a robust sense' (King, 2016, p. 54).

At the same time, it is obvious that Dana should not be blamed for using assistive technologies, and this is for two reasons. The first is that graphic organisers do require both some cognitive effort (or cognitive credit) and moto-perceptual skills to

be used, i.e. data have to be organised, constantly updated, and – most importantly – interpreted. The second reason why Dana should not be blamed is that cognitive credit, in the sense intended by K&V, is not the only variable that should be considered in the assessment of the capacities of cognitive tool users. Indeed, there are several cases in which the absence of cognitive credit has a positive valence, can be identified without attributing any negligence to subjects, but rather appears to be connected to the epistemic virtues of the subjects (Pritchard, 2013). Consider a plane pilot or a surgeon who perceptually accesses check lists to perform their tasks correctly and to reduce the risk of omissions. These agents transform a purely cognitive memory task (remembering the procedure) into a cognitive and perceptual one, by integrating the reading of the list in their working routine (Sterelny, 2004; Hutchins, 1999). Compare them with colleagues who are more self-confident, or simply lazy, and who rely on their memory. Of course, the latter deserve more cognitive credit, but in these cases that credit goes hand in hand with a higher risk of error (Gawande, 2009) and major responsibilities in the event of an accident.

Consider Emma, a colleague of Dana and Ava who suffers from LD but does not acknowledge her difficulties and tries to use only her neurons for her strategic planning and decision making. Emma expends a lot of mental energies on this attempt. Her cognitive credit is obviously higher than that of Ava and Dana, but unfortunately the results are pretty poor. Similarly, a patient suffering from a severe form of Alzheimer's could easily fall into error when she tries to solve a memory task by relying on her neurons and not on a cognitive aid, because in this way she does not acknowledge her condition and will often make mistakes. In all these cases, the effective exploitation of a cognitive artifact to support human performance goes at the expense of the cognitive credit, but constitutes a form of meta-cognition in the tool-use (Heersmink and Knight, 2018) that is connected to the recognition and acceptance of personal cognitive limitations. These agents perform a low degree of metacognitive credit, because they do not rely on the most reliable and efficient resource to complete the task at hand. Hence, on some occasions more cognitive credit can be attributed with a loss of cognitive performance, while, conversely, in others the absence of cognitive credit and the cognitive advantages of an artifact are connected to the employment of meta-cognitive capacities, something we can call meta-cognitive credit.

In our view these examples reveal two important aspects both of the EXT-MN defended by K&V and of EXT itself. The first is that the charge of K&V that EMB fails to properly attribute cognitive credit to LD subjects is weak and grounded on a simplistic dichotomy between the presence/absence of cognitive credit, whereas this can be attributed in varying degrees. That critique does not take into account other important capacities and the degree of meta-cognitive credit. Both Ava and Dana use assistive technologies. Their cognitive effort is different compared with someone who does not use artifacts, for instance Emma, but also compared with each other. At the same time, Ava and Dana demonstrate good meta-cognitive capacities that are recognised and valued by EMB. The deeper meaning of EMB, which in our view gets lost in K&V's argumentation, lies in recognising the value of the ability to exploit artifacts to accomplish tasks that would be much more difficult (or impossible) to accomplish by individual subjects (Hutchins, 1999). This ability involves several capacities that cannot be reduced simply to "cognitive credit".

The second point that we can infer from the above is that the EXT-MN fails to properly account for these cases, and this because once it postulates the presence of an extended system, the different relationships of the mind with the artifact disappear, because we have only one extended cognitive system postulated for a normative and/or moral reason. Ava and Dana, when using their different assistive technologies properly integrated in their cognitive system, are considered by the EXT-MN to have the same cognitive credit. In both cases, according to the EXT-MN it would be unfair to say that they are not able to perform those cognitive operations, but as far as Ava is concerned, this seems quite controversial. Similarly, it seems difficult to maintain that the A.I. software has improved her cognitive capabilities. The relationship between her mind and the artifact is quite different than that of Dana. The software completely substitutes her brain, and this has different ethical consequences. However, the EXT-MN seems to be a sort of black box in which, once the EXT system is postulated, these important aspects of the relationship between the subject's mind and the artifact are lost.[15]

To conclude our critique, let us return to the cognitive credit assumption and to PI. The former is implicitly assumed in K&V's formulation, while the latter is explicitly rejected by them (as a misleading assumption of EMB). However, here we argued that when the relationship between the artifact and the human mind is simple and straightforward, it appears difficult to deny that PI has an explanatory power and that it may be useful, at least in some cases (for instance for identifying cheating). This seems evident when we compare the case of Ava, in which the cognitive work is almost completely delegated to the artifact, with the case of Dana, who uses a graphic organiser with a bigger amount of cognitive credit. In our view, embracing EXT and considering both the graphic organiser and Chat-GPT as extensions of the mind would be to conceal some important differences between them. In other words, sometimes the difference of cognitive credit exists and it matters, and it would be misleading to ignore it. However, in our view, the bigger mistake of K&V is to embrace the cognitive credit assumption while overlooking the existence of other key elements that have to be taken into account in the assessment of the tool-users, namely moto-perceptual skills and the meta-cognitive credit. The examples of Emma and of plane pilots or surgeons show that sometimes the diminishment of the cognitive credit goes hand in hand with better cognitive performance, and this demonstrates that PI is not an absolute principle.

4 Final critiques and concluding remarks

As we have seen, the EXT-MN depends on the methodological assumption, according to which we should, on a descriptive level of analysis, embrace the view that proves to be more useful and preferable on an ethical and pragmatic level (Levy, 2007, p.

[15] More generally, we agree with Sterelny on this point, when he compares the scaffolded and extended frameworks, arguing that "the extended mind picture is not heuristically helpful. It obscures rather than highlights both the continuities and the differences amongst external resources and their contributions to cognitive competence" (Sterelny, 2010, p. 473).

220; Clowes, 2013, p. 129; King, 2016, p. 45; Vold, 2018, p. 497Farina and Lavazza, 2022a, p. 8; 2022b; 2022c, p. 11). This assumption seems working as the opposite of the naturalistic fallacy, which derives a normative stance from a mere a description of the state of affairs (Ridge, 2019). The methodological assumption instead states that we should choose the best description and explanation of a phenomenon by virtue of a normative and ethical considerations.[16] However, none of these authors appealing to this assumption have tried to justify it on theoretical grounds, but rather have simply used it as an ad-hoc solution for offering reasons to overcome the EMB/EXT impasse. Thus, we believe that this assumption generally needs further theoretical justification from the EXT-MN supporters.

Conversely, our standpoint is that in order to achieve the most socially and morally desirable effects, we should first embrace the most convincing and informative theory on a descriptive level of analysis. It is only thanks to a good previous description of the complex dynamics involved in agent-artifact cognitive interactions that we can appropriately assess normative trade-offs and principles.

However, K&V may defend the methodological assumption by grounding its legitimacy on ethical and political grounds. They may argue that the EXT-MN they propose is a form of 'ameliorative project' in favour of the social and political status and treatment of people suffering from cognitive disability. There are different kinds of 'ameliorative projects' meant to support different kinds of socially oppressed and marginalised groups. For instance, according to Howard and Aas, "in an ameliorative project, the explanatory power of different models of race and gender are partly evaluated by whether they can help us combat sexism and racism in our society" (Howard & Aas, 2018, p. 1116). Thus, it seems that K&V's AAC is meant to boost an ameliorative project for those who are suffering from marginalisation due to their cognitive disability, offering them the socially useful explanatory tool of the Extended Cognition Theory (EXT). There may be various important factors in evaluating the explanatory power of models of cognitive disability and cognitive assistance, however Howard and Aas argue that.

hermeneutic justice' is a central one: our concepts should help oppressed people make sense of how their diverse and varying experiences of oppression intersect. […] our basic question has to be: what account of disability can be best put to use for our social purposes?" (Howard & Aas, 2018, p. 1116).

Thus, K&V may argue that, in some circumstances, EXT offers the best explanatory model, since it helps to de-marginalise cognitively impaired individuals. However, even if the theoretical legitimacy of this kind of ameliorative project is accepted, we have shown that AAC tacitly depends upon the morally problematic assumption of an ableist conception of agency, which assumes an intrinsically unworthy picture of cognitive disability and cognitive impairment, going against the spirit of an ameliorative project in favour of cognitive disability (Howard & Aas, 2018). Thus, EXT,

[16] We are thankful to Lawrence Shapiro for having stressed this point at the conference celebrating the 25 years of "The Extended Mind" article, organized in Boulder in August 2023.

within the EXT-MN only superficially offers a more *inclusive* and *progressive*[17] picture of cognitive disability.

According to the ameliorative project for cognitive disability, we should choose our concepts also for their capacity to inform people of their condition, enabling hermeneutic justice (Howard & Aas, 2018, p. 1116). However, EXT in this case fails to support hermeneutic justice because of its uninformativeness. In fact, if we accept the EXT-MN and its assumption that every artifact that is deeply integrated improves subjects' cognitive capabilities "in a robust sense" (King, 2016, p. 54), we should conclude that their disabilities disappear thanks to the improvement provided by the artifact. When focusing only on the extended systems, there is nothing really relevant that differentiates an LD subject using assistive technology from a subject using her own mind. In this sense, EXT when used as K&V do, becomes a black box (see Sect. 2.2) that ultimately undermines the agents' ability to make sense of their specific individual and contextual conditions qua embedded cognitive agents' who rely in complex ways on cognitive artifacts. To be clear, in the literature there are many proponents of EXT that clearly define the necessary and/or sufficient conditions for cognitive extension, under which it is possible to evaluate whether specific cases of agent-artifact integrations are instances of EXT. The point is that K&V do not define clearly any necessary/sufficient condition for EXT. Rather, by using the EXT-MN they would consider as extended cognitive systems forms of agent-artifact integration that may be very different, such as Dana + graphic organiser and Ava + ChatGPT, hiding and obscuring what distinguishes them. Thus, since the AAC is based on the EXT-MN, it does not seem to us a better way to assess the capacities of LD individuals, but rather a way to trivialise them. In this way, the extended mind becomes an all-encompassing phenomenon in which the internal differences and the cognitive mechanics at play in the relationship between the mind and the artifact fade into the background. Once the mind has been considered automatically as extended, the artifact is no more an artifact in relationship with the human mind, but a simple part of the mind itself, as such completely integrated with the rest of the mind.

In addition, despite K&V referring to LD individuals' self-reports on using assisting tools, giving voice to the epistemic and social subjects involved, they do so only instrumentally to support the EXT-MN, by abstracting the individuals from their specific context and skills. Once we have abandoned the assumptions underlying the AAC, the fact that assisted individuals feel more independent and self-reliant thanks to the use of assisting devices does not directly imply that we should embrace EXT to properly account for those experiences on normative grounds. Thus, it seems to us that appealing to students' self-reports (Young & Specht, 2010; 2011) as well as the entire AAC are forms of philosophical abstraction and conceptual exploitation (or at least instrumental use) of cognitive disability (Carlson, 2009; Timpe, 2019, p. 160).

[17] In this treatment we appeal to the values of inclusiveness and progressiveness as intended in Cassinadri (2022); Farina and Lavazza (2022a). The former is defined as "vastness and a plurality of ends and ways of flourishing" and the "autonomy of people in their choices", while the latter is defined in terms of accepting positive change in new directions (Farina & Lavazza, 2022a, p. 14).

5 Conclusion

By assuming the presence of a metaphysical and explanatory stalemate between EMB and EXT, we argued that two normative arguments in favour of EXT are flawed. The argument from assessment of capacities (AAC) proposed by King (2016) and Vold (2016), as well as Clowes' (2013) cognitive diminishment argument (CDA), fail to present EXT as normatively and morally preferable to EMB. We have argued that both arguments assumed a flawed dichotomic way of thinking, according to which either we extend the cognitive agency and credit of tool-users by considering them as enhanced extended cognitive systems, or, by embracing EMB, we reduce and diminish their cognitive agency, capacities, and credit. Moreover, we argued that CDA is flawed and uninformative on a descriptive level of analysis, because it wrongly assumes that the EXT/EMB distinction straightforwardly overlaps with the cognitive enhancement/diminishment one.

In addition, the EXT-MN supported by AAC and CDA is also morally undesirable because it assumes an intrinsically marginalising picture of cognitive disability, inherent to an ableist conception of human agency, which implies that whenever an agent is dependent on an external resource, she is less worthy, independent, autonomous and interesting (Clowes, 2013, p. 128). Thus, the assumption that relational and dependent agents are necessarily more marginalised and diminished motivates King (2016), Vold (2018), Clowes (2013), and Farina and Lavazza's (2022a) moral urge to extend the boundaries of such agents by extending their cognitive system. We have argued that they have tried to overcome this assumed marginalisation in the wrong way, by extending cognition and the boundaries of the agential system, thus obscuring the presence of cognitive deficits. However, to properly foster an inclusive and demarginalising picture of cognitive disability, we need to recognise the presence of cognitive deficits instead of hiding them, so that the welfare state can adequately provide for them in social and educational contexts (Day & Edwards, 1996; Cassinadri, 2022, p. 15).

To conclude, we have highlighted that positing extended cognitive systems by using these two normative arguments leads to the impossibility of discrimination between different types of relationships between the mind and the artifact. Indeed, the EXT-MN promotes a picture of EXT that does not enable us to properly evaluate on normative grounds the morally relevant implications of specific kinds of cognitive hybridisation. By abandoning the underlying assumptions of AAC and CDA, we can distinguish, on a more fine-grained level of analysis, different cases of cognitive diminishment and cognitive enhancement, and acknowledge the proper level of credit due to both disabled and non-disabled individuals according to the specific context.

Acknowledgements We are deeply thankful to the reviewer, who helped us to make our arguments clearer and sharper as well as to the organizers and participants of the conference TXM@25 ("The Extended Mind" at 25, University of Colorado, Boulder), where we presented this work, for their insightful feedback and comments.

Funding Open access funding provided by Scuola Superiore Sant'Anna within the CRUI-CARE Agreement.

Declarations

Conflict of interest with respect to his authorship or the publication of this article.

References

Ahlvik-Harju, C. (2015). Disturbing bodies – reimagining comforting narratives of embodiment through Feminist Disability studies. *Scandinavian Journal of Disability Research*, *18*(3), 222–233. https://doi.org/10.1080/15.017419.2015.1063545.

Barker, M. J. (2010). From cognition's location to the epistemology of its nature. *Cognitive Systems Research*, *11*, 357–366.

Brownlee, K., & Cureton, A. (Eds.). (2009). *Disability and disadvantage*. Oxford Academic.

Campbell, & Fiona Kumari. (2009). *Contours of Ableism: The production of disability and abledness*. Palgrave Macmillan. https://doi.org/10.1057/9780230245181.

Carlson, L. (2009). Philosophers of intellectual disability: A taxonomy. *Metaphilosophy*, *40*, 552–566.

Carr, N. (2008). Is Google making us stupid? *Yearbook of the National Society for the Study of Education*, *107*(2), 89–94.

Carr, N. (2010). *The shallows: How the internet is changing the way we think, read and remember*. Atlantic Books.

Cassinadri, G. (2022). Moral reasons not to Posit Extended Cognitive systems: A reply to Farina and Lavazza. *Philosophy and Technology*, *35*, 64. https://doi.org/10.1007/s13347-022-00560-0.

Chouinard, V. (1997). Making space for disabling differences: Challenging Ablest geographies. *Environment and Planning D: Society & Space*, *15*, 379–387.

Clark, A. (2011). Finding the Mind. In Book symposium on *Supersizing the Mind: Embodiment, action, and cognitive extension* (2008). *Philosophical Studies* 152: 447–461.

Clark, A. (2022). Extending the predictive mind. *Australasian Journal of Philosophy*. 1–12.

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

Clowes, R. W. (2013). The cognitive integration of E-memory. *Review of Philosophy and Psychology*, *4*, 107–133.

Day, S. L., & Edwards, B. J. (1996). Assistive Technology for Postsecondary Students with Learning Disabilities. *Journal of Learning Disabilities*, *29*(5), 486–492. https://doi.org/10.1177/002221949602900503.

Fabry, R. E. (2018). Betwixt and between: The enculturated predictive processing approach to cognition. *Synthese*, *195*, 2483–2518.

Farina, M., & Lavazza, A. (2022a). Incorporation, transparency, and cognitive extension. Why the distinction between embedded or extended might be more important to ethics than to metaphysics. *Philosophy and Technology*, *35*, 10.

Farina, M., & Lavazza, A. (2022b). Why there are still Moral reasons to Prefer extended over embedded: A (short) reply to Cassinadri. *Philosophy and Technology*, *35*, 67.

Farina, M., & Lavazza, A. (2022c). Mind embedded or extended: Transhumanist and Posthumanist reflections in support of the extended mind thesis. *Synthese*, *200*, 507.

Fasoli, M. (2016). Neuroethics of cognitive artifacts. In A. Lavazza (Ed.), *Frontiers in Neuroethics* (pp. 67–82). Cambridge Scholars Press.

Fasoli, M. (2018). Substitutive, complementary and constitutive cognitive artifacts: Developing an Interaction-centered Approach. *Review of Philosophy and Psychology*, *9*, 671–687.

Francis, L. P., & Silvers, A. (2007). Liberalism and individually scripted ideas of the Good: Meeting the challenge of Dependent Agency. *Social Theory and Practice*, *33*, 311–334.
Garner, J. B., & Campbell, P. H. (1987). Technology for persons with severe disabilities: Practical and ethical considerations. *The Journal of Special Education*, *21*(3), 122–132.
Gawande, A. (2009). *The checklist manifesto: How to get things right*. Metropolitan Books.
Heersmink, R. (2012). Mind and artifact: a multidimensional matrix for exploring cognition- artifact relations. *Proceedings of the 5th AISB Symposium on Computing and Philosophy*, eds J. M. Bishop and Y. J. Erden (Birmingham: AISB): 54–61.
Heersmink, R. (2015). Dimensions of integration in embedded and extended Cognitive systems. *Phenomenology and the Cognitive Sciences*, *14*, 577–598.
Heersmink, R. (2017). Extended mind and cognitive enhancement: Moral aspects of cognitive artifacts. *Phenom Cogn Sci*, *16*, 17–32.
Heersmink, R., & Knight, S., (2018). Distributed learning: Educating and assessing extended cognitive systems. *Philosophical Psychology*, *31*(6), 969–990. https://doi.org/10.1080/09515089.2018.1469122
Howard, D., & Aas, S. (2018). On valuing impairment. *Philosophical Studies*, *175*, 1113–1133. https://doi.org/10.1007/s11098-018-1074-y.
Hutchins, E. (1995). *Cognition in the Wild*. MIT Press.
Hutchins, E. (1999). Cognitive artifacts. In R. A. Wilson, & F. C. Keil (Eds.), *The MIT encyclopaedia of the cognitive sciences* (pp. 126–128). MIT Press.
Kersten, L. A. (2022). New Mark of the Cognitive? Predictive Processing and Extended Cognition. *Synthese*, *200*, 281.
King, C., & Kuhn, D. (20162000). Learning disability and the extended mind. *Essays in PhilosophyCurrent Directions in Psychological Science*, *179*(25), 38–68. https://doi.org/10.1111/1467-8721.00088.
Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science*, *9*(5), 178–181. https://doi.org/10.1111/1467-8721.00088
Levy, N., & MacIntyre, A. (20071999). *Neuroethics: Challenges for the 21th century*. Cambridge University PressOpen Court Publishing Co.
Mackenzie, C., & Stoljar, N. (2000). *Relational autonomy: Feminist perspectives on autonomy, agency, and the social self*. Oxford University Press.
Marconi, D. (2005). Contro La Mente Estesa. *Sistemi Intelligenti*, *17*(3), 389–398.
Menary, R. (2007). *Cognitive Integration: Mind and Cognition Unbounded*, Palgrave Macmillan.
Menary, R. (2018). Cognitive Integration How Culture Transforms Us and Extends Our Cognitive Capabilities. In Gallagher, S., Newen, A., & Bruin, L De. (Eds.), *Oxford Handbook of 4E Cognition* (pp. 187–215). Oxford University press.
Metcalfe, J., & Shimamura, A. P. (Eds.). (1996). *Metacognition: Knowing about knowing*. MIT Press.
Pritchard, D. (2013). Epistemic virtue and the epistemology of education. *Journal of Philosophy of Education*, *47*(2), 236–247.
Reader, S. (2007). The other side of agency. *Philosophy*, *82*(4), 579–604.
Ridge, M. (2019). Moral Non-Naturalism. The Stanford Encyclopedia of Philosophy, Edward N. Zalta (ed.) https://plato.stanford.edu/archives/fall2019/entries/moral-non-naturalism/.
Ross, D. (2010). and James Ladyman The Alleged Coupling-Constitution Fallacy and the Mature Sciences. In Menary (Ed), *The Extended Mind* (pp. 154 – 66).
Rowlands, M. (2010). *The new science of the mind: From extended mind to embodied phenomenology*. MIT Press.
Rowlands, M. (2010a). *The new science of the mind: From extended mind to embodied phenomenology*. MIT Press.
Rupert, R. (2009). *Cognitive systems and the extended mind*. Oxford University Press.
Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, *333*(6043): 776–778.
Sprevak, M. (2010). Inference to the hypothesis of extended cognition. *Studies in History and Philosophy of Science*, *41*, 353–362.
Sterelny, K. (2004). Externalism, Epistemic artefacts and the extended mind. In R. Schantz (Ed.), *The Externalist Challenge* (pp. 239–255). Walter de Gruyter.
Sterelny, K. (2010). Minds: Extended or scaffolded? *Phenomenology of the Cognitive Sciences*, *9*, 465–481.
Tarvainen, M. (2019). Ableism and the life stories of people with disabilities. *Scandinavian Journal of Disability Research*, *21*(1), 291–299. https://doi.org/10.16993/sjdr.632.

Timpe, K. (2019). Moral ecology, disabilities and human agency. *Res Philosophica*, 96(1): 17–41. Timpe, K. (2022). Cognitive Disabilities, Forms of Exclusion, and the Ethics of Social Interactions. *The Journal of Philosophy of Disability*, 2: 157–184.

Vold, K. (2018). Overcoming deadlock: Scientific and ethical reasons to embrace the extended mind thesis. *Philosophy and Society*, 29(4): 471–646.

Wasserman, D., Asch, A., Blustein, J., & Putnam, D. (2017). Cognitive Disability and Moral Status. The Stanford Encyclopedia of Philosophy, Edward N. Zalta (ed.), https://plato.stanford.edu/archives/fall2017/entries/cognitive-disability/.

Wheeler, M. (2016). A tale of two dilemmas: Cognitive kinds and the extended mind. In C. Kendig (Ed.), *Natural kinds and classification in scientific practice* (pp. 210–221). Routledge.

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). Oxford University Press.

Wolf, M., & Stoodley, C. (2008). *Proust and the Squid: The Story and Science of reading brain*. Icon Books.

Woollett, K., & Maguire, E. A. (2011). Acquiring the knowledge of London's layout drives structural brain changes. *Current Biology*, *21*(24), 2109–2114.

Young, G., & Jacqueline Specht. (2011). *Attending a Demonstration School: Its Impact on Students' Self-Concept and School Motivation*. University of Western Ontario.

Young, G., & Specht, J. (2010). Assistive technology: Its impact on self-concept and motivation across school transitions. Presented at the *Bridges to Learning Conference*, Mississauga, ON.

Authors and Affiliations

Guido Cassinadri[1] · Marco Fasoli[2]

✉ Guido Cassinadri
guido.cassinadri@santannapisa.it

Marco Fasoli
marco.fasoli@uniroma1.it

[1] Scuola Superiore Sant'Anna, Piazza Martiri della Libertà, 33, Pisa 56127, Italy

[2] Dipartimento di Filosofia, Università di Roma La Sapienza, Via Carlo Fea, 2, Roma, RM 00161, Italy