Science & Education
https://doi.org/10.1007/s11191-025-00660-1

ARTICLE

Generative Artificial Intelligence and Extended Cognition in Science Learning Contexts

Angel Rivera-Novoa[1] · **Daniel Augusto Duarte Arias**[1,2]

Received: 9 September 2024 / Revised: 17 May 2025 / Accepted: 23 May 2025

Abstract

This paper philosophically examines the impact of generative artificial intelligence on learning processes from the perspective of extended cognition. The central problem addressed is how these technologies can transform students into passive or active learners, influencing the development of cognitive skills. It will be argued that generative artificial intelligence presents risks of diminishing cognitive activity among students, as it is likely to substitute—rather than complement—the cognitive subject. It will also be argued that there are ways to leverage generative artificial intelligence so that learners are not passive but rather active cognitive subjects. Three cases will be presented, with empirical support, to show how this leveraging is possible: the production of feedback, assistive technologies, and gamification. In these cases, generative artificial intelligence is a complementary cognitive artifact rather than a substitutive one. To achieve this goal, the paper presents the framework of the extended mind thesis as a conducive scenario for analyzing the relationship between generative artificial intelligence and learning contexts, and it analyzes specific cases of science education. An analysis of the types of cognitive artifacts will also be conducted to examine how generative artificial intelligence intervenes in learning in both substitute and complementary ways.

Keywords Extended cognition · Generative artificial intelligence · Cognitive artifact · Learning · Educational technology

1 Introduction

The presence of information and communication technologies in educational contexts is increasingly noticeable and has grown significantly since the confinement due to the COVID-19 pandemic (Kang, 2021; Kashif & Shujjaudin, 2023). Indeed, education

✉ Angel Rivera-Novoa
angel.riveran@udea.edu.co

Daniel Augusto Duarte Arias
daduartea@usbcali.edu.co

1 Instituto de Filosofía, Universidad de Antioquia, Medellín, Colombia

2 Universidad de San Buenaventura, Cali, Colombia

detached from the use of information and communication technologies would seem to be less and less realistic (Valverde-Berrocoso et al., 2021). In addition, the recent rise of generative artificial intelligence has made available a series of tools that produce texts, images, videos, and other types of content that were often requested as outputs of formative processes (García-Méndez et al., 2024; Olga et al., 2023). This proliferation of artificial intelligence tools undoubtedly raises the question of the appropriateness of their integration in learning contexts, in any knowledge area, including science education.

While some defend the use of these artificial intelligence tools to enhance student learning (Kadaruddin, 2023; Lee et al., 2024), others seem to question their use by virtue of some consequences on cognitive and behavioral development, or even because of several possible ethical and academic integrity issues (Skulmowski, 2024; Ye et al., 2024). Recent discussions about the presence of artificial intelligence in learning contexts, therefore, seem to oscillate between a marked optimism leading to staunch advocacy, on the one hand, and a strong pessimism leading to an outright rejection of its implementation, on the other. This oscillation is present in discussions on science education, as the use of generative artificial intelligence in this field is a fact (Wang et al., 2024), and while some concerns about excessive dependence on these tools have arisen (Tang & Cooper, 2024; Wells, 2024), it has also been argued that the use of these new technologies has potential in science education (Heidt, 2024; Lin et al., 2024).

The question of the consequences of the use of this type of technological tools on learning processes is interwoven with the reflection that can be made about the relationship between human cognition and technology. Hence, an approach such as that of extended cognition (Clark, 2008; Clark & Chalmers, 1998; Pritchard, 2010; Rowlands, 2010) is attractive for addressing the issue of the use of artificial intelligence for learning purposes. According to this approach, external tools and objects may not only be a means to accomplish cognitive tasks but may be a constitutive or integral part of cognition. In the scientific context, this approach could have merit to be explored, as in the scientific enterprise the use of diagrams, maps, equations, simulations, and other external cognitive artifacts to represent natural phenomena is common (Tang, 2024). Hence, the integration of generative artificial intelligence brings both challenges and potential beneficial uses.

The aim of this paper is to analyze conceptually, within the philosophical framework of the extended cognition paradigm, the impact of the use of artificial intelligence in learning contexts, with a special emphasis on science education. The thesis we wish to defend is that the use of generative artificial intelligence runs the risk of making learners passive subjects in learning processes, which could eventually be detrimental. However, it will also be argued that there are ways to leverage generative artificial intelligence so that learners are not passive but rather active cognitive subjects. Hence, a better approach to understanding the role of these tools in learning contexts should involve analyzing specific extensions of our cognition in this kind of artificial intelligence.

To achieve our goal, we will first present the general framework of what is known as the extended cognition or the extended mind and its relation to artificial intelligence (Section 2). Next, we will show how generative artificial intelligence runs the risk of taking away all epistemic credit from learning subjects in the performance of cognitive tasks, which would mean that their use in learning contexts could become detrimental (Section 3). Then, we will present three cases that show how generative artificial intelligence can function as a complementary artifact in learning contexts and science education by emphasizing an active role for learners: the production of feedback, assistive technologies, and gamification (Section 4). We will end the article with some brief conclusions (Section 5).

2 The Extended Mind and the Enhancement of Learning

Traditional cognitive sciences argue that cognitive processes occur in the brain. Some classical approaches, such as Jerry Fodor's, argue that the mind is a sequential data processing system, similar to the functioning of computational systems (Fodor, 1975). Other approaches, such as connectionism, claim that the mind functions as a network linking information processing and the perceptual system (Hatfield, 2014). In contrast to these traditional views, the situated mind approach emerges based on the theses of Kirsh and Maglio (1994) and Hutchins (2006), who explain that there is a constitutive, and not only causal, link between the mind, the body, and the world.

The extended mind thesis has arisen in this situated mind framework. In particular, Clark and Chalmers (1998) argue that mental processes do not all occur exclusively in the brain but may be extended outside in an interaction with the world. Thus, advocates of extended cognition oppose the intracranial view by arguing that it is possible to make use of external tools for the accomplishment of tasks involving cognitive processes. In other words, tools, the environment, and the body can configure a cognitive system.

The debates surrounding this extended mind thesis have raised different waves that try to explain how this intimate relation of the mind with external elements is possible. The first wave argues that cognitive processes can be performed physically in different external elements because some of the cognitive processes of external structures can function as if they were occurring inside the brain. This thesis is known as the parity principle. For example, activities that require memory could resort to external tools such as notebooks, digital devices, or agendas to store information that would later be available for use. Thus, these external elements would play an active role in cognitive processes that extend outside the skull due to their functional similarity (Clark & Chalmers, 1998; Wheeler, 2010).

The second wave of the extended mind abandons the functionalist explanation to argue that the mind is extended because external elements complement cognitive processes. The agent and the environment may have different properties and functionalities, but both contribute in a complementary way within a system by coupling for the accomplishment of a cognitive task. Consequently, the agent, when manipulating external elements, integrates to different degrees, resulting in different couplings where internal and external processes occur (Menary, 2010; Sutton, 2010).

The third wave of the extended mind suggests that the mind is characterized by the dynamism of cognitive processes, its flexible boundaries that are constantly open to negotiation, and the distribution of the cognitive assemblage, which as a whole depends on cultural, social, and material aspects. This wave emphasizes that the mind consists of heterogeneous elements and, therefore, interactions with the environment vary according to the given contexts. Thus, the degrees of integration will not depend exclusively on the agents, but also on the environments in which the agents develop their cognitive activities or tasks. Furthermore, this wave argues that consciousness has a fundamental role in the extension of the mind (Kirchhoff & Kiverstein, 2019; Sterelny, 2010).

If the thesis of the extended mind is true, there would be optimistic scenarios in which human beings can enhance themselves cognitively. For some authors, this enhancement would lead to *Homo technologicus* (Duarte Arias & Ortega Chacón, 2024). Other arguments, much more critical, suggest that cognitive hybridization with external elements, as the extended mind thesis states, is what human beings have always naturally done (Rivera Novoa, 2020). Nevertheless, the extended mind thesis offers some characteristics such as transparency, niche construction, and plasticity, from which it is worth examining the integration between external elements and the human being.

When extended mind theorists refer to transparency, they argue that agents commonly resort to tools for the accomplishment of cognitive tasks. However, these agents are only aware of the use of these tools when events such as tool damage or loss occur. This happens because the agent focuses its attention on the task and not on the tool being used to achieve the goal. In this way, transparency would explain in which contexts it can be stated that there is or is not an extended mind due to the interaction of the agent with the near environment (Andrada, 2020).

Niche construction is an explanatory model through which some authors try to show how the mind adapts to different contexts, just like human biological organs. In these cases, the mind would have an arrangement similar to that of organs, such as those that contribute to digestion (for example, in the use of fire for the cooking and digestion of food prior to consumption). In the same way, the mind resorts to external scaffolding to support and enhance its cognitive processes (Sterelny, 2010).

The disposition of the mind, as shown by the construction of niches and transparency, is some of the cases that would explain the plasticity of the mind. Plasticity presents itself in different degrees, and each one of them depends on how much the mind is integrated or not with external elements. In this sense, the human being can achieve a deep embodiment with external elements, which would guarantee a bodily configuration in which internal and external processes are linked in the same system. In this way, the mind adapts to bodily and environmental structures to provide the most appropriate solution to the tasks facing an agent (Clark, 2007).

One clear scenario in which an extended cognitive process can take place is when an agent tries to solve a mathematical operation or a scientific problem, such as balancing a chemical equation or calculating the trajectory of a moving object, which is relevant for science education through external tools. Sometimes, agents resort to their internal processes to perform various operations, such as calculations or mathematical functions using memory, reasoning, abstraction, etc. In the scientific laboratory, this process is even more evident: researchers use pipettes, balances, spectrometers, and specialized software that not only facilitate their work but also fundamentally transform how they think about the phenomena being studied. While agents can perform these tasks by making use of their internal cognitive processes, according to extended mind theorists, they can also do so by resorting to tools such as a notebook and pencil, a calculator, an abacus, or a smartphone. In this sense, if external tools play an active role in the achievement of these cognitive tasks, and their processes occur as if they were carried out in the head, then there would be extended cognitive processes and integration with external tools.

In science education and related fields, a variety of artificial intelligence tools already serve as cognitive aids that extend learners' abilities. For example, in laboratory courses, AI-driven data analysis software can automatically perform tasks like curve-fitting on experimental data or run complex simulations, helping scientists and students interpret results more efficiently (Zielesny et al., 2011). Intelligent tutoring systems (ITS) in subjects such as physics or mathematics act as adaptive mentors that provide personalized feedback and hints, guiding learners through problem-solving steps while offloading some routine computations (Shafiq et al., 2025). Similarly, agent-based modeling programs enable biology students to explore complex ecosystems by simulating the interactions of numerous agents (predators, prey, etc.), allowing learners to observe emergent patterns without manually computing each interaction (Dickes & Sengupta, 2013). Advanced machine-learning techniques are also making their way into the classroom; for instance, neural networks have been used to analyze genomic datasets, uncovering patterns in DNA or protein data that would be infeasible to find by hand (Novakovsky et al., 2023). Even widely available symbolic computation tools (e.g., Wolfram Alpha) can instantly solve calculus problems, serving as external cognitive tools that students can leverage in learning (Dimiceli et al., 2010). These integrations would also lead to an enhancement in the cognitive process because these machines perform information processing much faster and operate with larger amounts of information. If artificial intelligence tools are part of a cognitive system, then it should be assumed that such a system is superior, to some degree, than one that makes use of paper and pencil or one that is only constituted by the internal processes in the brain to develop cognitive tasks.

The extended cognition thesis has allowed different authors to explain how the link with tools such as artificial intelligence, assistive technologies, or other types of tools is possible and how these integrations, in some contexts, support and assist educational processes. In science education, microscopic observation, manipulation of laboratory equipment, or interpreting complex visualizations can present significant barriers for students with various disabilities. Pritchard et al. (2021) are optimistic that education can make use of assistive technologies to meet the developmental needs of children and youth. The authors argue that if assistive technologies can be integrated into special educational settings, they can produce extended cognitive integration. Pritchard et al. (2021) defend the use of technologies in education, questioning those contexts in which the use of these tools is not justified by the functioning of traditional education:

> shouldn't there be a standing obligation in even a mainstream educational setting, where feasible, to provide technological/environmental solutions, ideally those that could become part of the student's extended cognition? Education is plausibly concerned with the enhancement of cognitive character, after all, so if there are ways of developing cognitive character which are specifically technological, then what reason, beyond mere tradition, is there for not exploiting them? (Pritchard et al., 2021, p. 19)

The proposal of these authors is that, as long as there are contexts in which technological and environmental tools can be used to enhance students' cognition, the educational environment is obliged to make use of these tools because educational processes are constantly concerned with cognitive enhancement. This leads to questions about some current educational scenarios that, during the COVID-19 pandemic, made use of technological tools to provide solutions to the atypical educational environment, but, subsequent to the isolation, discontinued these practices by abandoning these methods of developing cognitive skills. Another possible scenario to examine, due to advances in recent years, is the use of generative artificial intelligence in the educational context and how it influences the development of cognitive skills or, on the contrary, whether it undermines certain skills acquired in traditional or conventional settings. That will be our topic in the next sections.

3 Generative Artificial Intelligence as a Substitutive Artifact in Learning Contexts

As pointed out by Aagaard (2021) and Bruineberg and Fabry (2022), there appears to be a "harmony bias" in studies of extended cognition. That is, the theoretical analyses and conceptual applications of the extended mind and extended cognition thesis focus mostly on cases where the extension of the mind has positive cognitive repercussions or, in general, where the integration with external elements is always successful. This leaves aside analyses where the technological extension of the mind results in some sort of cognitive detriment or, in general, in consequences that we would not characterize as "good" for cognitive subjects. These scenarios could, moreover, cover broad swaths of our mental lives. Bruineberg and Fabry (2022) point out, for example, that the habitual and diversionary use of our smartphones is a case of extended mind-wandering and covers a large spectrum of our actual mental activity.

The harmony bias in studies of extended cognition has thus privileged the analysis of cases such as belief formation or memory enhancement through cognitive extensions on the Web (Heersmink & Sutton, 2020; Smart, 2017), as well as dimensions such as affectivity (Colombetti & Roberts, 2015), diagnoses and treatments in psychiatry (Hoffman, 2016), and, of course, educational contexts of teaching and learning (Pritchard, 2016; Pritchard et al., 2021). In all these cases, for example, there is a harmony between external technological tools and the biological mind, such that the cooperation between the two poles configures or constitutes a unified cognitive circuit. An emerging area of study now relates the extended cognition thesis to artificial intelligence. However, once again, there is an emphasis on how the integration of artificial intelligence technologies enhances human cognition. The essential idea is that, just like any artifact, artificial intelligence can become part of our cognitive circuits and, moreover, can enhance them (Helliwell, 2019; Nyholm, 2024).

In recent years, a special type of artificial intelligence has attracted attention: generative artificial intelligence. These are intelligent models based on advanced deep learning algorithms capable of creating new content such as texts, images, music, videos, and other sets of content that were normally produced by humans. Generative artificial intelligence can have significant impacts on several dimensions of human life. For example, thanks to this type of intelligence, particularly through generative adversarial networks with reinforcement learning, the chemical design of new drugs has been accomplished by generating new molecules and properties (Vanhaelen et al., 2020). Nonetheless, this kind of intelligence has gained greater popularity with the launch of chatbots that allow lay users to give simple instructions to generate content. Such intelligences include ChatGPT, Claude, Midjourney, and DALL-E. Of course, having tools such as ChatGPT, which is capable of creating texts as well written as (or better than) those of humans, raises questions about the convenience of integrating such intelligence into learning scenarios. Although there are optimistic views about their use in educational contexts given that such tools can create customized content designed for particular needs and contexts (Kadaruddin, 2023; Lee et al., 2024; Lin et al., 2024), there is also the suspicion that an over-reliance on such tools could undermine the cultivation of skills and the attainment of knowledge by learners (Tang & Cooper, 2024; Wells, 2024). Nyholm, though not referring specifically to educational contexts but to the cultivation of intelligence, expresses this concern in the following words:

> A key question here, however, is whether delegating to AI technologies the tasks for which we use our intelligence could potentially be a way of making us less intelligent. [...] If we hand over too many tasks to AI systems, and we therefore have fewer

occasions or incentives to develop our capacity for intelligence, then there is a risk that we might end up being less intelligent than we could otherwise be. (Nyholm, 2024, p. 80)

The problem lies in the fact that using generative artificial intelligence in learning contexts may prevent students from learning to write texts, analyze the writings of others, make inferences, interrelate ideas, make deductions, and, in general, carry out the necessary activities that typically lead to the formation of critical thinking and the development of cognitive and epistemic skills. Within the framework of the extended mind, the question is whether the extension of our cognitive abilities and skills into generative artificial intelligence truly favors learning or, on the contrary, constitutes an impediment to it. If the latter, we would be facing a scenario where cognitive extension into these tools constitutes non-harmonic cognitive circuits—the type of cognitive extensions that, as we have mentioned, have been overlooked in the literature due to the harmony bias. In the remainder of this section, we will show that the integration of generative artificial intelligence in educational contexts *may* not favor learning processes.

Duncan Pritchard, one of the most notable current proponents of the extended cognition thesis, presents an argument for the inclusion of technologies in learning contexts, without specifically referring to generative artificial intelligence. In his text "Intellectual Virtue, Extended Cognition, and the Epistemology of Education" (2016), Pritchard examines the impact of technologies in educational contexts and their potential to undermine students' abilities to perform specific tasks. Pritchard introduces a key distinction between two epistemological approaches. On the one hand, we have "epistemic individualism," which does not consider that cognition can be extended but rather holds that it is exclusively instantiated in biological individuals and conceives technological tools merely as auxiliary means. On the other hand, we have "epistemic anti-individualism" which adopts an extended cognition perspective where cognitive processes then go beyond the skull and the skin. According to Pritchard, it is only under the lens of epistemic individualism that concerns about the loss of cognitive skills due to dependence on technological tools emerge (Pritchard, 2016, pp. 119, 122, 125). Indeed, once the extended cognition thesis is accepted, the use of technologies for learning processes, including generative artificial intelligence, would in no way undermine cognition but rather extend it. For that reason, fears of undermining learning would have no place in the framework of the extended mind thesis.

Our approach to the problem, in Pritchard's terms, is "anti-individualistic." To that extent, we argue, along with proponents of the first and second waves of the extended mind thesis, that cognition is a phenomenon that does not take place exclusively in a biological individual but can be coupled or supplemented by external tools and artifacts, including generative artificial intelligence, in learning contexts. However, we argue that not every extension of the mind is equivalent to cognitive enhancement. In other words, there *may* be cognitive extensions that lead to detriments in learning processes. We also believe that generative artificial intelligence *may* be an instance of this type of situation. Hence, we believe that it is essential to analyze the types of artifacts into which the mind can be extended in order to examine whether *some* cognitive extensions through generative artificial intelligence *can be* considered detrimental to learning contexts.

Fasoli draws a distinction between three types of cognitive artifacts. Cognitive artifacts are "physical objects that have been created or modified to contribute to the completion of a cognitive task" (Fasoli, 2018, p. 681). A cognitive task, in turn, simply is the activity, structured in terms of means and ends, that leads to the attainment of a primary cognitive goal, which is usually the attainment of knowledge. Thus, finding a location is a cognitive task. A map or a GPS would be cognitive artifacts insofar as they contribute to the attainment or performance of the task. This definition of cognitive artifact aligns with the thesis of cognitive extension since artifacts can constitute cognitive circuits together with the subject that uses them. Now, the three types of cognitive artifacts proposed by Fasoli are complementary, substitutive, and constitutive. The complementary cognitive artifact is one that assists in a process that could exist independently. For instance, the use of pencil and paper to perform arithmetic operations. The process could occur purely internal to the subject, independently of the existence of such elements. Substitutive cognitive artifacts, on the other hand, assume the cognitive work necessary to complete a task, such as a GPS navigation system that replaces the need for orientation. Finally, constitutive cognitive artifacts are those that are necessary for a task. For example, in the task of reading, the presence of a text is necessary (Fasoli, 2018, pp. 678–680).

It is important to highlight that the same artifact can sometimes be a substitute, sometimes complementary, and sometimes constitutive. Hence, we are cautious in the formulation of the thesis we are defending, according to which the use of generative artificial intelligence *runs the risk* of making the learners passive subjects, thereby impairing their learning. This is quite different from saying that the use of these tools *necessarily* implies a detriment in learning contexts. For this reason, use is a fundamental aspect of our analysis. The tool or artifact, by itself, cannot dictate whether it is beneficial or detrimental in educational settings. Nonetheless, we do not want to argue that tools and artifacts are totally neutral either. The point is that the analysis, if it is really based on the extended cognition thesis and an anti-individualistic epistemic approach, must take use and tool as an indivisible unit of study. Indeed, if we argue that we can extend our mind and cognition into generative artificial intelligence, and we intend to examine whether such extension is beneficial for learning, it would be a mistake to focus exclusively on artifacts to address the issue. It would also be wrong to focus only on the use that is made of the cognitive artifact. The use and the artifact constitute the cognitive process as a whole.

The same cognitive artifact, then, can be complementary, substitutive, or constitutive depending on its use. The camera of our smartphone, for example, can help us remember the details of a moment or place, in which case it would be complementary. But it can also be substitutive if we are unable to remember anything in particular on our own. Moreover, it can also be constitutive if our task is to analyze the photograph itself or to read something written in the image. We will focus our attention on complementary and substitutive artifacts, leaving aside the case of constitutive artifacts.[1] Now, let us consider generative artificial intelligence. Through a simple instruction, a chatbot is able to write a text on artificial intelligence and learning. The chatbot can follow our indications, for example, regarding the length or even the language of writing. However, the same chatbot, with a different yet equally simple instruction, could, instead of producing the text, provide a

[1] The constitutive aspect of cognitive artifacts is not in the same as the constitutive aspect of the extended cognition thesis. The former establishes that we cannot do certain cognitive task without the presence of a specific artifact—for instance, we cannot read a text without the presence of the text itself. The latter refers to the idea that a subject and an object, artifact, or tool can form a coupled system. Accordingly, we could use a constitutive cognitive artifact without extending our cognition in that tool (Cassinadri, 2024, p. 13). In the reading activity, we truly need the text to perform the task. Nonetheless, we are not a coupled system with the text. On the contrary, we can solve mathematical problems with or without a tool, and when we use some tool, we may constitute a coupled system with it. Hence, we will focus on the constitutive aspect of the extended mind thesis, rather than the constitutive aspect of certain cognitive artifacts.

brainstorm that offers a clearer idea of how the text might be written, for example, as a kind of peer that helps generate knowledge, as seen in Oh and Lee (2024). In the first case, the chatbot is a substitutive artifact, since it is replacing the subject's activity almost entirely. In the second case, on the other hand, the chatbot is a complementary cognitive artifact. Here, the activity of writing the text remains in the hands and control of the individual. The chatbot has merely provided a series of clues about how to structure the writing. In other words, the subject has not been replaced by the tool.

We can argue that both substitutive and complementary artifacts are susceptible to cognitive extension. However, in the case of substitution, the learner's activity is almost nil. This is not the case with complementarity, where the learner has simply made use of tools to generate certain ideas. We can state that, in the first case, the learner did not write the text, while in the second case, the learner did. The issue with generative artificial intelligence tools is that, as they can function as both substitutive and complementary cognitive artifacts, they can have a negative impact on learning processes when their integration in educational contexts is completely substitutive. Therefore, even if we accept that the use of generative artificial intelligence can constitute a case of cognitive extension in learning contexts, their integration into these contexts *runs the risk* of turning learners into passive individuals, when this kind of intelligence functions as a substitute cognitive artifact. This would undoubtedly negatively affect the learning processes. Therefore, the substitutive use of artificial intelligence in learning contexts may configure a case of non-harmonious cognitive extension.

In virtue of the short lifespan of ChatGPT and other similar tools, research on this issue is scarce. However, there are analyses on how this substitution could occur. Ye et al. (2024) provide empirical support for the thesis that the use of ChatGPT fosters "inert thinking," understood within Kahneman (2011) dual-process framework, which conceptualizes cognition as the interplay between an intuitive, unreflective system and a more conscious, analytical one. Inert thinking is associated with the former, and reliance on chatbots such as ChatGPT appears to encourage this passive mode of thought while inhibiting active cognition. Moreover, the positive responses to chatbots, stemming from their perceived effectiveness, further perpetuate this dynamic. These findings support the claim that employing generative artificial intelligence in learning contexts risks a substitutive effect, rendering learners passive recipients rather than active participants in their educational processes.

In Rivera-Novoa (2024), it is philosophically argued that a reliance on certain technologies that substitute human cognitive activity can lead to a distinct type of ignorance. This ignorance is not defined by a lack of propositional knowledge; rather, it involves the notion that by extending tasks to technology in a substitutive manner, we forfeit the experience of autonomous thinking. Central to this argument is the notion of "cognitive phenomenology," referring to one's awareness of what it is like to be someone that thinks, remembers, calculates, and so on. If generative artificial intelligence is allowed not merely to complement but to substitute tasks that require active engagement, we risk remaining ignorant of this cognitive phenomenology—an outcome symptomatic of learners becoming passive rather than active subjects.

In science education, the use of generative artificial intelligence can be substitutive as well. Wang et al. (2024) analyze the impact of this intelligence on problem-solving in STEM education. Through a survey with college students in the USA, Wang et al. (2024) show that generative artificial intelligence is used to interpret results, explore related topics, or summarize papers. Nonetheless, over half of the students reported that they simply ask a chatbot to solve the problem, and 38% of students simply copy and paste the problem that they should try to solve (Wang et al., 2024, p. 11). Although the majority of students reported that they are conscious of the risk of using generative artificial intelligence, as this tool is prone to misinformation, producing nonsensical solutions to the problems, and overreliance could affect their learning (Wang et al., 2024, pp. 13–14), the problem-solving ability in science education could be substituted by the tool. In a meta-analysis of studies on the use of generative artificial intelligence in science education, Tang (2024) makes some recommendations and, among them, explicitly points out that when interacting with this tool, one should try to obtain “outputs to complement, not replace, students’ transduction across different representations in interactive ways” (Tang, 2024, p. 1348). It is clear that the nature of this type of intelligence can tend to replace rather than complement cognitive activity in science education.

The fact that the extension of our cognition to generative artificial intelligence tools risks being replaced does not only imply the abandonment of the implementation of our epistemic skills; it also entails other consequences that we can qualify as undesirable. For example, Skulmowski (2024) empirically shows how overdependence on chatbots creates the illusion that we have better skills than we really have; that is, a placebo effect is produced. The illusion of having acquired or exercised one’s own skill arises when, in fact, the task has been done by the chatbot. This is accompanied by the “ghostwriter effect,” which consists of attributing to oneself the authorship of content when, in reality, the task has been done by the tool. This is without mentioning the most obvious problems, such as the risk of inheriting the biases inherent in these technologies (Cooper & Tang, 2024; Wells, 2024) or the risk of taking bad information as correct—or, as they are now called, “hallucinations” (Kadaruddin, 2023)—within which various types can be found: overfitting, logic and reasoning errors, mathematical errors, factual errors, and so on (Sun et al., 2024). The substitution of the student’s cognitive activity by generative artificial intelligence is especially problematic in science education, where materiality and direct experience with natural phenomena are fundamental for the construction of scientific knowledge, but are lost with the continuous interaction with the tool (Tang & Cooper, 2024). This further results in an overvaluation of the tool as an epistemic authority (Cooper, 2023), leading to results that are unreliable, lacking in scientific novelty and irreproducible by virtue of LLM hallucinations (Wells, 2024), as well as biased representations of the scientific activity itself—for example, in the production of stereotypical images of scientific environments (Cooper & Tang, 2024).

4 Generative Artificial Intelligence as a Complementary Cognitive Artifact

The very nature of generative artificial intelligence makes it typically a substitutive cognitive artifact. To the extent that generative artificial intelligence is capable of autonomously creating content, its users become passive subjects in performing cognitive tasks. Nonetheless, is it possible that cognitive extension in generative artificial intelligence results in a use of it that is complementary rather than substitutive?

Fasoli (2018) exemplifies how the use of a GPS is typically substitutive but could become complementary. GPSs are tools that perform the task of indicating the most suitable route to reach a destination. According to Fasoli, an individual can use a GPS without making use of his or her own sense of location at all. This would be a case where the device is a substitute, as we can do without our ability to orient ourselves in an unknown space. However, the same GPS can be complementary if, for example, a tourist previously looks at a map to try to find his destination, then walks without using the GPS, and only uses it to confirm some detail during his journey. In such a case, the individual forms a coupled system with the GPS without the orientation activity being performed by the device. Thus, although GPS can *typically* be substitutive, we can make complementary use of it.

Could we say the same for generative artificial intelligence? In the previous section, we suggested that a brainstorm produced by a Chatbot could be a case of complementary extension. Cassinadri (2024) has argued that ChatGPT can be used in a complementary way for concept apprehension, critical thinking training, and the cultivation of intellectual virtues. But the cases of the brainstorm and Cassinadri's proposals still need empirical studies. In this section, we will present three empirically supported cases, where there is cognitive extension in generative artificial intelligence and where it becomes a complementary artifact in learning contexts. The three situations we will analyze are (i) the use of generative artificial intelligence for feedback, (ii) the use of such intelligences as assistive technologies, and (iii) the integration of complementary gamification environments with generative artificial intelligence.

4.1 Multi-agent Feedback and Socratic LLM

It is true that chatbots are *typically* substitutive, as they are content producers. However, learning contexts can integrate them as complementary cognitive artifacts in a manner analogous to some uses of GPS. This is the case of *feedback* production. As Lim et al. (2023) and Oh and Lee (2024) point out, the use of chatbots such as ChatGPT for text production has led to these tools replacing human writing in learning contexts. Hence, it is thought that it is necessary to eliminate writing assessments (Zhai, 2022), as it has even been shown that these tools can approve law exams (Kelly, 2023) or medical licensing exams (Hammer, 2023). However, generative artificial intelligence can be used to identify conceptual gaps in learners (Lim et al., 2023), improve critical thinking skills (Dickerson, 2025), or generate feedback for student performance (Guo et al., 2024; Lim et al., 2023; Namoun et al., 2024; Wongvorachan et al., 2022).

Feedback is beneficial to a learner to the extent that it provides information that can bridge the gap between the learner's actual performance and the desired or imagined performance. Feedback gives the learner the ability to identify where they are in the learning process, what mistakes they are making, and how they can try to overcome them. With the emergence of generative artificial intelligence, teachers can generate personalized feedback for each learner by pinpointing instructions for their performance (Lim et al., 2023; Guo et al., 2024). In addition, learners themselves can generate such reports on their own. As Namoun et al. (2024) point out in their review, this has been shown to benefit the understanding of complex notions, the learning of other languages, the improvement of writing skills, the test preparation, and the improvement of learner engagement. In science education, personalized feedback is a useful tool for the development of inquiry skills. Lin et al. (2024) implemented the GPT-Assisted Summarization Aid (GASA), a tool that integrates generative artificial intelligence to provide formative feedback at various stages of experiential learning in STEM. Unlike traditional systems, GASA can analyze students' scientific explanations and offer suggestions that promote deeper, evidence-based reasoning, without replacing the learner's cognitive activity.

However, as Steiss et al. (2024) and Jansen et al. (2024) point out, generative artificial intelligence feedback often suffers from over-praise and over-inference. Over-praise is the situation in which the learner's performance is overestimated through the evaluation, giving the learner a wrong picture of his or her actual performance. On the other hand, over-inference occurs when the feedback given does not depend only on the learner's contributions, but on inferences that go beyond and that are made by artificial intelligence. Both situations distort the real performance situation of the learner, preventing him from taking measures that lead to the improvement of his learning.

Guo et al. (2024) developed a multi-agent model with generative artificial intelligence that produces higher quality feedback than those delivered by a single agent (e.g., a single chatbot), precisely reducing over-praise and over-inference. Such a model, which they call "Autofeedback" was developed to provide feedback on assessments written by science students. The multi-agency nature of the Guo et al. (2024) proposal lies in the fact that each of these intelligent agents is dedicated to certain specific designated tasks. The model works roughly as follows. The student performs the delivery. A first chatbot, following precise instructions, performs a feedback report. Then, a second agent (another chatbot) examines, reviews, and if necessary, corrects the first feedback. In this case, the instruction should specifically ask for the decrease of over-praise and over-inference (Guo et al., 2024, pp. 4–6). This results in more reliable feedback and may result in better learning on the part of students. Indeed, the feedback generated with the multi-agent model reduced the occurrence of over-praise by 14.17%, of over-inference by 20.21% relative to the reports generated by a single agent (Guo et al., 2024, p. 11).

In this case, the use of generative artificial intelligence turns out to be complementary. In no case is a chatbot replacing a learner's writing activity. While the feedback is produced by generative artificial intelligence, it can be used by the learners to improve their performance, their evaluations, and, in general, their understanding of a particular subject. Artificial intelligence cooperates with the learner to perform a cognitive task. Hence, the generated feedback reports constitute a case where there is an extension of non-detrimental cognition. In these cases, the tool is not a substitute for our own cognition. The technological extension, in this case, does not produce writing or work directly but facilitates our capacity for self-criticism, which would be more cognitively demanding without the feedback.

Dickerson (2025) presents a feedback model in which a chatbot is programmed to have linguistic exchanges in a Socratic style. In this case, the chatbot is configured to ask specific Socratic questions to the learner (e.g., "what is justice?", "what is courage?"), and based on the answers given, the chatbot questions the definition given and asks a counter-question. In response to the new answers, the chatbot continues to criticize the answers and asks again, demanding a refinement of the analysis and argumentation. Tools such as these, which are easy to set up, can contribute to the development of critical thinking, academic linguistic exchange, and argumentation in general.

In tune with the Socratic model, Tang (2024, pp. 1340–1341) considers that the nature of generative artificial intelligence promotes its dialogical interaction, which results in effective use of this tool in science education. This effective use, according to Tang (2024), is mediated by the fact that the student should not confuse this type of tool with search engines such as Google, so that they avoid considering the tool as a scientific epistemic authority (Cooper, 2023). In contrast, the outputs should be critically evaluated. One way to achieve this is to encourage the use of chatbots to formulate dialogical questions for understanding scientific concepts. Tang suggests that LLMs can help generate spaces with diverse perspectives, so that uncritical dependence on the tool fades and students develop skills of their own.

The case of the Socratic chatbot developed by Dickerson (2025) and the use of LLM proposed by Tang (2024) can be understood as a kind of Socratic feedback. And like the multi-agent model, the Socratic interaction with the chatbot can be understood as a coupled, extended system that can eventually lead to epistemic aims. Thus, tools can be developed with generative artificial intelligence, in which there is an extension that is not substitutive but complementary, such that learning is enhanced. Hence, although artificial intelligence is *typically* a substitutive cognitive artifact, like GPS, it can be used in a complementary way—also like GPS—making the learners active subjects in their process.

4.2 Generative Artificial Intelligence as an Assistive Technology

Assistive technologies are devices that offer a better quality of life to individuals with various types of disabilities (Preum et al., 2021). Some of these devices are part of cognitive orthotics, which are platforms intended "to support learning, memory, keeping records, making documents and organizing the thoughts" (Zdravkova, 2022, p. 253). Hence, their linkage to learning environments allows complementing the cognitive processes of the tasks developed by students with disabilities.

Generative artificial intelligence can be integrated into these tools, since due to its modeling it promises to improve service, accessibility, and adaptive content to the functional diversities of subjects (Smith et al., 2023). This type of intelligence has the potential, due to its multiple applications and its ability to generate diverse content, to complement and support processes in users with cognitive deficits or disabilities, by offering unique elements such as personalization by adjusting to training data of the same individuals (Griffith & Rathore, 2023). Another useful resource of this type of intelligence is modeling real-time faces that contribute to communication and emotional contact for autistic individuals through deepfake tools (cf. Giri & Brady, 2023). In such cases, generative artificial intelligence is assistive because it contributes to cognitive processes due to the support, improvement, or enhancement of tasks performed by users with disabilities.

Heidt (2024) shows how generative artificial intelligence is transforming accessibility in science education for people with various neurodivergent conditions or chronic illnesses. Individuals with severe brain injuries, paralysis, or conditions such as autism use these tools to organize scientific tasks, prioritize actions, retrieve specialized information, and write scientific papers (Heidt, 2024, p. 462). Heidt highlights the use of image generators for people with aphantasia—that is, people without visual imagery—enabling them to understand visually complex scientific concepts and speech-to-text transcribers for students with mobility impairments (Heidt, 2024, p. 463). Tang (2024) points out precisely how the representation of scientific concepts is very important in science education and how it can be empowered by generative artificial intelligence tools that produce images and representations. In the case of people with aphantasia, as Heidt (2024) points out, this is particularly important, as it promotes inclusion and equity in science education. Hence, Heidt (2024) points out that policies prohibiting the use of these tools in science education contexts should be avoided, as such prohibition would be contrary to the inclusion of minority groups.[2]

[2] Although the use of generative artificial intelligence for translation is not exactly a case of assistive technology, it is a use that can promote inclusion by helping to bridge language barriers in science. Generative artificial intelligence tools—such as Google Translate or DeepL—can support researchers for whom English is a second language, enabling them to access, produce, and communicate scientific knowledge more effectively (Tang, 2024). These tools assist with tasks such as reading and understanding academic literature and improving the structure and clarity of written communications—from research papers to emails. By reducing the extra time and effort these scientists often must spend on language-related tasks and enhancing the quality of their work, generative AI helps to *level the field* for non-native English speakers in academia (Heidt, 2024).

In arguing for the extended mind thesis, Clark and Chalmers (1998, pp. 12–15) present a case that has become paradigmatic: the case of Otto. Such a case can illustrate how these assistive technologies supported by generative artificial intelligence can be a case of complementary cognitive extension. Otto is a man suffering from Alzheimer's disease, and, consequently, his biological memory is deficient. Otto writes down relevant information in his notebook, which acts as his "external memory." For Otto, all he needs to do is consult this device to retrieve what he needs to remember. For Clark and Chalmers, Otto's notebook "is central to his actions in all sorts of contexts, in the way that an ordinary memory is central in an ordinary life" (1998, p. 13). As we can see, a rudimentary assistive technology, such as Otto's notebook, is a clear example of extended cognition for Clark and Chalmers. For this reason, assistive technologies with generative artificial intelligence should also constitute cases of extended cognition, insofar as their use allows the attainment of epistemic aims such as remembering, in Otto's case.

How do these assistive technologies with generative artificial intelligence come into play in learning contexts? Several recent reviews draw attention to their potential in such scenarios and the increased research interest in their application for educational purposes (Fu et al., 2025; Mustafa et al., 2024; Tlili et al., 2024). UNESCO (2023) also highlights this potential to assist students with special needs, such as those with visual or hearing limitations. Although this potential has been identified, only four countries (China, Jordan, Malaysia, and Qatar) had reported in 2023 that their governments recommend the use of generative artificial intelligence to assist students with needs and achieve more inclusive learning contexts (UNESCO, 2023, p. 37).

In Almufareh et al. (2024), it is shown how natural language processing (NLP) tools generate inclusive learning scenarios for students with speech and hearing impairments. Indeed, these tools can generate written texts in real time, which facilitates comprehension and communication for learners with such problems, ensuring access to information that would otherwise be more difficult to obtain. In Yang et al. (2024), it is shown how generative artificial intelligence can be used for visually impaired people. There, a model (VIAssist) is developed in which, through the uploading of photographs, detailed descriptions of the photographs are obtained. In cases where the photographs are not well registered, which is very common in people with these impairments, the model suggests taking them again, indicating the correct way to obtain the necessary information for the description. Tools of this type facilitate the acquisition of knowledge for students with visual limitations, allowing the creation of more inclusive learning environments.

As in these cases, students with dyslexia can benefit from the design of tools with generative artificial intelligence to improve their learning. Recent studies show that systems such as AI4LA (D'Urso & Sciarrone, 2024), TutorChat (De Marco et al., 2024), or Karaton (Hauwaert et al., 2020), designed with generative artificial intelligence, provide personalized support by analyzing conversational data and generating concept maps, facilitating the visual representation of knowledge and understanding of academic content, and also training users in distinguishing the type of words they most often confuse through algorithmic data analysis. Preliminary evaluations of these tools in educational settings have shown positive results in supporting reading comprehension, synthesis, and information search tasks, traditionally challenging areas for students with dyslexia. Despite these positive results, as shown by Dabaghi et al. (2024), developments of these tools have been very scarce.

As in the case of Otto's notebook, students with some kind of hearing impairment, visual impairment, aphantasia, or dyslexia can use assistive technologies with generative artificial intelligence to achieve epistemic purposes in learning contexts. And like Otto's notebook, such tools can be part of coupled cognitive circuits, so their use can be considered cognitive extension. As Pritchard et al. (2021) rightly point out about the use of assistive technologies, “here is not merely a cognitive process that employs AT (i.e., as subject-and-instrument-a non-systems approach), but rather an extended cognitive process (and thus extended cognition) that has AT as a proper part” (p. 15). In addition, the assistive use of these technologies makes them complementary cognitive artifacts because they support individuals in their training and learning to develop tasks without in any way substituting them. In effect, the users of these technologies continue to play an active role in their learning processes, only now with a support that facilitates access to and understanding of information.

4.3 Gamification and Generative Artificial Intelligence

Gamification is the use of game elements and technological environments in non-game contexts. Some elements that characterize this type of technological tool include badges, scoreboards, levels, and avatars. The purpose of these resources is to stimulate the learner to achieve greater focus, whether through persistence or repetition in achieving objectives. Moreover, gamification employs collaborative participation or competition among different learners (Buckley & Doyle, 2017; Ding, 2019). Sailer and Homner (2020) state that the effects of gamification can be significant for learning, as some subsets of the empirical analysis show optimal indicators in competition and collaboration, although gamification based on collaboration yields better results in terms of achievement of knowledge.

Now, as Andrade et al. (2016) and Prieto-Andreu (2024) point out, the negative effects of gamification on learning have been largely ignored in the literature. For instance, Boulet (2012) and Kshetri (2024) are critical of gamification implementation. Boulet (2012) argues that while some learning materials can benefit from elements of game mechanics, gamification could be overused, leading to a waning of interest in the tool over time. Kshetri (2024), in turn, holds that the use of gamification could lead to an increase in cheating in learning contexts, as competitive settings stimulate these practices.

Despite the misgivings against gamification, it can be argued that, at first glance, such tools are complementary cognitive artifacts, rather than substitutes. Indeed, gamification requires the active participation of its users. Even in cases where gamification challenges loss of interest and avoidance of cheating, its effective uses are not possible without the active participation of the learner.

Now, what about generative artificial intelligence in relation to gamification? It is important to start by clarifying two things. Firstly, gamification and generative artificial intelligence are matters of different natures. While the former can be considered an “environment” that attempts to enhance learning, the latter is a specific type of artificial intelligence that aims at content creation. Secondly, the relationship between them is only just being explored and still moves in the realm of speculation. According to the review by Abbes et al (2024), between 2019 and 2024, only four research papers were found that explicitly linked gamification with generative artificial intelligence. Furthermore, Huber et al. (2024) have argued that the interaction between gamification and LLM is exclusively limited to the creation of games and educational materials.

In the remainder of this section, we will show that, despite the scarcity of research, which may be due to the novelty of generative artificial intelligence, there are some studies, beyond the four just mentioned, that show that there are good reasons to think that gamification with this type of intelligence can be understood as a sort of complementary and not substitutive cognitive extension. Nonetheless, the use of generative artificial intelligence in gamified scenarios must go beyond the issue of the design of such scenarios or educational materials. The key to the issue, we will argue, lies in the power of generative artificial intelligence to create novel interactions and the personalization of content, because it is thanks to this that there is a complementarity between learners and this type of intelligence. It is not a matter of learners obtaining desired content through prompts. Instead, generative artificial intelligence should be able to offload them from various tasks so that they can obtain epistemic aims that would otherwise be more difficult to obtain.

In addressing the issue of assistive technologies, we have brought the case of Karaton (Hauwaert et al., 2020). This tool, which aims to help students with dyslexia improve their reading comprehension, is a gamification setting. The tool offers a set of mini-games that are highly personalized. Its algorithms, through data provided by users, identify error patterns (e.g., confusing "b" with "d"). In addition, the algorithm outputs a series of exercises in the form of new mini-games focused on working on the identified patterns. Hauwaert et al., (2020, p. 94) point out that the results in reading improvement are at least as effective as traditional methods, although engagement and motivation are enhanced by virtue of personalization and the playful environment. In this sense, Karaton is a case of extended cognition, where a coupled system is configured, integrating the personalization produced by artificial intelligence to achieve cognitive aims. Here, the users play an active role, so that in no case does the tool substitutes the exercise of the learner, but complements it, favoring the identification of their difficulties and overcoming them.

Zhang et al. (2024) developed a gamified tool with generative artificial intelligence to combat the effects of bubble filters on information consumption. Generative artificial intelligence allows the creation of five characters with completely different personalities. Users can interact with them through various conversations. The tool included two gamified features to incentivize interaction with multiple perspectives: the "Viewpoints Puzzle" and an assessment task with multiple-choice questions. By completing conversations with all agents and answering questions correctly, users were able to "illuminate" pieces of a puzzle, gaining a sense of accomplishment and a more complete understanding of the topic. The study showed positive results in engagement with the understanding of others' points of view. As mentioned earlier, Tang (2024) shows the importance of this dialogical perspective in education. Indeed, the linguistic potential of generative artificial intelligence can be harnessed for the formation of critical thinking by discussing the nature of various scientific concepts. Multi-agent models, which can be built with this type of intelligence, promote the development of critical understanding of scientific concepts. This configures a case of extended cognition, where the interaction between the user and the gamified system with generative artificial intelligence produces a coupling that enhances information processing and critical thinking skills. As in the previous case, artificial intelligence does not replace the user in any way, but rather the user maintains a constant active role. The tool complements the user's activity to facilitate the understanding of different points of view.

In a similar vein, but with virtual reality immersion, Song et al. (2024) developed LearningverseVR. This is an immersive multiplayer platform that leverages generative artificial intelligence, such as Zhang et al. (2024), to create characters and to engage them in dialogues through LLM. Song et al. (2024) point out that some advantages of this platform are the personalization of learning, which is gained by interacting with the characters created with artificial intelligence; the possibility of situated and experiential learning, which is enabled by the simulation of situations; and the motivation generated by gamified elements such as rankings and medals, which are also the product of generative artificial intelligence and depend on each situation. The platform is not designed for specific learning but can be used for a multitude of purposes, such as learning languages, performing science education simulations, learning about history, and so on. However, Suraj Kumar et al. (2024) describe how interactive simulations of climate systems allow climate science students to manipulate variables and observe the consequences in real time, fostering a deeper understanding of complex causal relationships in climate science. As in the case of Zhang et al. (2024), the platform developed by Song et al. (2024) and the studies by Suraj Kumar et al. (2024) could be considered instances of extended cognition in generative artificial intelligence, insofar as the interaction with the characters and the environments created by them can configure an integrated cognitive circuit with the user. Moreover, as in the two previous cases, the artificial intelligence is not substituting any cognitive activity of the learner, but rather complementing it.

Thus, the production of feedback, assistive technologies, and gamification scenarios, in conjunction with generative artificial intelligence, are cases of extended cognition. In these cases, several epistemic aims, such as self-criticism, identifying and overcoming difficulties or limitations, and critical thinking are complemented by interaction with generative artificial intelligence tools. And although, by its nature, this kind of intelligence is *typically* substitutive, in the cases we have presented, it becomes a complementary cognitive artifact, since it neither allows the agents' passivity nor becomes a surrogate for cognition.

5 Conclusions

In this paper, we philosophically examine the role of certain applications of generative artificial intelligence in learning contexts, taking science education as the main focus of analysis, without reducing the reach of our argument to it. From the perspective of extended cognition, we have analyzed how this type of artificial intelligence can typically affect students' active participation in their educational process, which could lead to an excessive dependence that would compromise the development of cognitive skills. These technologies can displace active learning, making it more passive if not properly managed. This risk arises because generative artificial intelligence is prone to become a substitute, rather than a complementary, cognitive artifact, resulting in cognitive tasks being performed almost entirely by the tool.

Conversely, we also discuss the potential benefits of generative artificial intelligence in several scenarios. The production of feedback, as in the multi-agent model of Guo et al. (2024) or in the Socratic use of LLMs (Dickerson, 2025), provides examples that show how learners can identify gaps in their learning processes and enhance critical abilities. In the case of assistive technologies, generative artificial intelligence can help users—e.g., blind, mute, or deaf students—access information that would otherwise be more difficult to obtain. Moreover, some tools could be developed to identify difficulties more effectively and overcome them, as in the case of dyslexia analyzed in Section 4.2 and Section 4.3. Finally, we have shown how several gamification scenarios can be designed with generative artificial intelligence as environments that improve critical thinking and the ability to understand different points of view (Zhang et al., 2024), assist impairments such as dyslexia (Hauwaert et al., 2020), or, in general, create favorable environments that enhance engagement and leverage generative artificial intelligence to foster interactivity and personalization in learning. While generative artificial intelligence is typically a substitutive cognitive artifact, these examples demonstrate cases of cognitive extension in which this kind of artificial intelligence becomes a complementary cognitive artifact. In such cases, learners behave in active rather than passive ways. In all these scenarios, learners interact with generative artificial intelligence to offload various processes so that they can achieve epistemic aims that would otherwise be more difficult to reach.

These three cases raise some practical considerations at the moment of application to avoid generative artificial intelligence becoming a substitutive cognitive artifact. When using generative artificial intelligence for feedback, generated content is not a replacement for learners' tasks but a tool that may enhance their processes. For assistive technologies, personalization enabled by this type of intelligence supports students with disabilities while keeping them actively engaged. Finally, in gamification, generative artificial intelligence should be used to enable novel interactions and personalized content to complement learners' own efforts, as opposed to simply automating content generation, which risks fostering a passive attitude. Across all these applications, the key is to leverage generative artificial intelligence capabilities in a targeted way to assist and complement human learning activities. These applications imply that learners and educators know the potential of generative artificial intelligence, which must be encouraged by institutions.

Future lines of research, therefore, should further explore how generative artificial intelligence can create a hybrid learning environment that does not displace human activity. In particular, it would be very useful to advance research that empirically and conceptually analyzes how to design learning environments where cognitive extension into generative artificial intelligence functions effectively as a complementary and non-substitutive artifact, ensuring that human activity remains a central element of this technological extension of the mind.

Acknowledgements We are grateful for the comments of Leandro Giri, Juan Toro, and Henry Camilo Bejarano on previous versions of this manuscript. We would also like to thank the reviewers of the paper for their good considerations to improve it.

Author Contribution Both authors contributed equally to the development of this paper. The conceptualization, research, and writing of the manuscript were the result of a collaborative effort and shared responsibilities.

Funding Open Access funding provided by Colombia Consortium. This paper is a product of the research project "Transhumanism and the extended mind thesis: Can technology radically transform human cognition?" with code 2023–59530 funded by the Vice-Rectory of Research of the University of Antioquia.

Data Availability This study is of a theoretical/conceptual nature and does not use empirical data. Therefore, no data availability statement is required.

Declarations

Ethics Approval Not applicable.

Conflict of interest The authors declare that they have no conflict of interest.

References

Aagaard, J. (2021). 4E cognition and the dogma of harmony. *Philosophical Psychology, 34*(2), 165–181. https://doi.org/10.1080/09515089.2020.1845640

Abbes, F., Bennani, S., & Maalel, A. (2024). Generative AI and gamification for personalized learning: Literature review and future challenges. *SN Computer Science, 5*(8), 1154. https://doi.org/10.1007/s42979-024-03491-z

Almufareh, M. F., Kausar, S., Humayun, M., & Tehsin, S. (2024). A conceptual model for inclusive technology: Advancing disability inclusion through artificial intelligence. *Journal of Disability Research*, 3(1). https://doi.org/10.57197/JDR-2023-0060

Andrada, G. (2020). Transparency and the phenomenology of extended cognition. *Límite: Revista de filosofía y psicología, 20*(15), 1–17.

Andrade, F. R. H., Mizoguchi, R., spsampsps Isotani, S. (2016). The bright and dark sides of gamification. In A. Micarelli, J. Stamper, spsampsps K. Panourgia (Eds.), *Intelligent Tutoring Systems* (9684, 176–186). Springer International Publishing. https://doi.org/10.1007/978-3-319-39583-8_17

Boulet, G. (2012). Gamification: The latest buzzword and the next fad. *eLearn, 2012*(12), 2407138.2421596. https://doi.org/10.1145/2407138.2421596

Bruineberg, J., & Fabry, R. (2022). Extended mind-wandering. *Philosophy and the Mind Sciences*, *3*. https://doi.org/10.33735/phimisci.2022.9190

Buckley, P., & Doyle, E. (2017). Individualising gamification: An investigation of the impact of learning styles and personality traits on the efficacy of gamification using a prediction market. *Computers & Education, 106*, 43–55. https://doi.org/10.1016/j.compedu.2016.11.009

Cassinadri, G. (2024). ChatGPT and the technology-education tension: Applying contextual virtue epistemology to a cognitive artifact. *Philosophy & Technology, 37*(1), 14. https://doi.org/10.1007/s13347-024-00701-7

Clark, A. (2007). Re-inventing ourselves: The plasticity of embodiment, sensing, and mind. *Journal of Medicine and Philosophy, 32*(3), 263–282. https://doi.org/10.1080/03605310701397024

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

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

Colombetti, G., & Roberts, T. (2015). Extending the extended mind: The case for extended affectivity. *Philosophical Studies, 172*(5), 1243–1263. https://doi.org/10.1007/s11098-014-0347-3

Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology, 32*(3), 444–452. https://doi.org/10.1007/s10956-023-10039-y

Cooper, G., & Tang, K.-S. (2024). Pixels and pedagogy: Examining science education imagery by generative artificial intelligence. *Journal of Science Education and Technology, 33*(4), 556–568. https://doi.org/10.1007/s10956-024-10104-0

D'Urso, S., spsampsps Sciarrone, F. (2024). AI4LA: An intelligent Chatbot for supporting students with dyslexia, based on generative AI. In A. Sifaleras spsampsps F. Lin (Eds.), *Generative Intelligence and Intelligent Tutoring Systems* (369–377). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-63028-6_31

Dabaghi, K., D'Urso, S., spsampsps Sciarrone, F. (2024). Artificial intelligence and learning of students with dyslexia: A brief review. In Z. Kubincová, T. Hao, N. Capuano, M. Temperini, S. Ge, Y. Mu, P. Fantozzi, spsampsps J. Yang (Eds.), *Emerging Technologies for Education* (155–169). Springer Nature. https://doi.org/10.1007/978-981-97-4243-1_13

De Marco, V., Sciarrone, F., & Temperini, M. (2024). TutorChat: A chatbot for the support to dyslexic learner's activity through generative AI. In *2024 IEEE International Conference on Advanced Learning Technologies (ICALT)* (155–157). IEEE. https://doi.org/10.1109/ICALT61570.2024.00051

Dickerson, P. (2025). Learning with Socrates: How generative AI and ancient pedagogy can develop students' critical thinking skills. In H. Crompton & D. Burke (Eds.), *Artificial Intelligence Applications in Higher Education* (pp. 90–105). Routledge.

Dickes, A. C., & Sengupta, P. (2013). Learning natural selection in 4th grade with multi-agent-based computational models. *Research in Science Education, 43*(3), 921–953. https://doi.org/10.1007/s11165-012-9293-2

Dimiceli, V. E., Lang, A. S. I. D., & Locke, L. (2010). Teaching calculus with Wolfram|Alpha. *International Journal of Mathematical Education in Science and Technology, 41*(8), 1061–1071. https://doi.org/10.1080/0020739X.2010.493241

Ding, L. (2019). Applying gamifications to asynchronous online discussions: A mixed methods study. *Computers in Human Behavior, 91*, 1–11. https://doi.org/10.1016/j.chb.2018.09.022

Duarte Arias, D. A., & Ortega Chacón, O. (2024). Ya no somos Homo Sapiens: Exploración a los desafíos del Homo Tecnologicus. *Tecné, Episteme y Didaxis: TED, 55*, 279–295.

Fasoli, M. (2018). Substitutive, complementary and constitutive cognitive artifacts: Developing an interaction-centered approach. *Review of Philosophy and Psychology, 9*(3), 671–687. https://doi.org/10.1007/s13164-017-0363-2

Fodor, J. A. (1975). *The language of thought*. Crowell.

Fu, B., Hadid, A., & Damer, N. (2025). Generative AI in the context of assistive technologies: Trends, limitations and future directions. *Image and Vision Computing, 154*, 105347. https://doi.org/10.1016/j.imavis.2024.105347

García-Méndez, S., de Arriba-Pérez, F., & Somoza-López, M. D. C. (2024). A review on the use of large language models as virtual tutors. *Science & Education*, 1–16. https://doi.org/10.1007/s11191-024-00530-2

Giri, D., & Brady, E. (2023). Exploring outlooks towards generative AI-based assistive technologies for people with Autism (Versión 1). *arXiv.preprint* arXiv:2305.09815. https://doi.org/10.48550/ARXIV.2305.09815

Griffith, H., & Rathore, H. (2023). Personalized aging-in-place support through fine-tuning of generative AI models. *Eighth International Conference on Mobile and Secure Services (MobiSecServ), 2023*, 1–2.

Guo, S., Latif, E., Zhou, Y., Huang, X., & Zhai, X. (2024). Using generative AI and multi-agents to provide automatic feedback. *arXiv preprint* arXiv:2411.07407. *https://doi.org/10.48550/arXiv.2411.07407*

Hammer, A. (2023). The rise of the machines? ChatGPT CAN pass US medical licensing exam and the bar, experts warn-after the AI chatbot received B grade on Wharton MBA paper. *Daily Mail.* https://www.dailymail.co.uk/news/article-11666429/ChatGPT-pass-United-States-Medical-Licensing-Exam-Bar-Exam.html

Hatfield, G. (2014). Cognition. In L. A. Shapiro (Ed.), *The Routledge handbook of embodied cognition* (pp. 361–373). Routledge.

Hauwaert, H., Ghesquière, P., Tordoir, J., spsampsps Thomson, J. (2020). Karaton: An example of AI integration within a literacy app. In K. Miesenberger, R. Manduchi, M. Covarrubias Rodriguez, spsampsps P. Peňáz (Eds.), *Computers helping people with special needs* (91–96). Springer International Publishing. https://doi.org/10.1007/978-3-030-58796-3_12

Heersmink, R., & Sutton, J. (2020). Cognition and the web: Extended, Transactive, or Scaffolded? *Erkenntnis, 85*(1), 139–164. https://doi.org/10.1007/s10670-018-0022-8

Heidt, A. (2024). 'Without these tools, I'd be lost': How generative AI aids in accessibility. *Nature, 628*(8007), 462–463. https://doi.org/10.1038/d41586-024-01003-w

Helliwell, A. (2019). Can AI mind be extended? *Evental Aesthetics, 8*, 91–119.

Hoffman, G. A. (2016). Out of our skulls: How the extended mind thesis can extend psychiatry. *Philosophical Psychology, 29*(8), 1160–1174. https://doi.org/10.1080/09515089.2016.1236369

Huber, S. E., Kiili, K., Nebel, S., Ryan, R. M., Sailer, M., & Ninaus, M. (2024). Leveraging the potential of large language models in education through playful and game-based learning. *Educational Psychology Review, 36*(1), 25. https://doi.org/10.1007/s10648-024-09868-z

Hutchins, E. (2006). *Cognition in the wild*. MIT Press.

Jansen, T., Höft, L., Bahr, L., Fleckenstein, J., Möller, J., Köller, O., & Meyer, J. (2024). Comparing generative ai and expert feedback to students' writing: Insights from student teachers. *Psychologie in Erziehung und Unterricht, 71*(2), 80–92. https://doi.org/10.2378/peu2024.art08d

Kadaruddin, K. (2023). Empowering education through generative AI: Innovative instructional strategies for tomorrow's learners. *International Journal of Business, Law, and Education, 4*(2), 618–625. https://doi.org/10.56442/ijble.v4i2.215

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kang, B. (2021). How the COVID-19 pandemic is reshaping the education service. In J. Lee spsampsps S. H. Han (Eds.), *The future of service post-COVID-19 pandemic, volume 1* (15–36). Springer Singapore. https://doi.org/10.1007/978-981-33-4126-5_2

Kashif, M., & Shujjaudin, M. (2023). Covid-19 is a catalyst for the emergence of information technology in education institutes. *Journal of Development and Social Sciences*, *4*(II). https://doi.org/10.47205/jdss.2023(4-II)44

Kelly, S. M. (2023). ChatGPT passes exams from law and business schools. *CNN Business, 26.* https://edition.cnn.com/2023/01/26/tech/chatgpt-passes-exams/index.html

Kirchhoff, M. D., & Kiverstein, J. (2019). *Extended consciousness and predictive processing: A third-wave view (1 [edition])*. Routledge.

Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science, 18*(4), 513–549. https://doi.org/10.1207/s15516709cog1804_1

Kshetri, N. (2024). The academic industry's response to generative artificial intelligence: An institutional analysis of large language models. *Telecommunications Policy, 48*(5), 102760. https://doi.org/10.1016/j.telpol.2024.102760

Lee, G. G., Mun, S., Shin, M. K., & Zhai, X. (2024). Collaborative learning with artificial intelligence speakers. *Science & Education*, 1–29. https://doi.org/10.1007/s11191-024-00526-y

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education, 21*(2), 100790. https://doi.org/10.1016/j.ijme.2023.100790

Lin, C.-J., Lee, H.-Y., Wang, W.-S., Huang, Y.-M., & Wu, T.-T. (2024). Enhancing reflective thinking in STEM education through experiential learning: The role of generative AI as a learning aid. *Education and Information Technologies*. https://doi.org/10.1007/s10639-024-13072-5

Menary, R. (2010). Cognitive integration and the extended mind. In R. Menary (Ed.), *The extended mind* (pp. 227–244). The MIT press.

Mustafa, M. Y., Tlili, A., Lampropoulos, G., Huang, R., Jandrić, P., Zhao, J., Salha, S., Xu, L., Panda, S., Kinshuk, López-Pernas, S., & Saqr, M. (2024). A systematic review of literature reviews on artificial intelligence in education (AIED): A roadmap to a future research agenda. *Smart Learning Environments, 11*(1), 59. https://doi.org/10.1186/s40561-024-00350-5

Namoun, A., Ibrahim, I. A., Mustafa, E., Alrehaili, A., Tufail, A., Shuja, J., Bilal, K., et al. (2024). Generative artificial intelligence in education: an umbrella review of applications and challenges. https://doi.org/10.21203/rs.3.rs-4892155/v2

Novakovsky, G., Fornes, O., Saraswat, M., Mostafavi, S., & Wasserman, W. W. (2023). ExplaiNN: Interpretable and transparent neural networks for genomics. *Genome Biology, 24*(1), 154. https://doi.org/10.1186/s13059-023-02985-y

Nyholm, S. (2024). Artificial intelligence and human enhancement: Can AI technologies make us more (artificially) intelligent? *Cambridge Quarterly of Healthcare Ethics, 33*(1), 76–88. https://doi.org/10.1017/S0963180123000464

Oh, P. S., & Lee, G. G. (2024). Confronting imminent challenges in humane epistemic agency in science education: An interview with ChatGPT. *Science & Education*, 1–27. https://doi.org/10.1007/s11191-024-00515-1

Olga, A., Tzirides, Saini, A., Zapata, G., Searsmith, D., Cope, B., Kalantzis, M., Castro, V., Kourkoulou, T., Jones, J., da Silva, R. A., Whiting, J., & Kastania, N. P. (2023). Generative AI: Implications and applications for education. *ArXiv*. https://doi.org/10.48550/ARXIV.2305.07605

Preum, S. M., Munir, S., Ma, M., Yasar, M. S., Stone, D. J., Williams, R., Alemzadeh, H., & Stankovic, J. A. (2021). A review of cognitive assistants for healthcare: Trends, prospects, and future directions. *ACM Computing Surveys, 53*(6), 1–37. https://doi.org/10.1145/3419368

Prieto-Andreu, J. M. (2024). Cómo Evitar Efectos Negativos al Gamificar en Educación: Revisión Panorámica y Aproximación Heurística hacia un Modelo Instruccional. *Multidisciplinary Journal of Educational Research, 14*(2), 1–23. https://doi.org/10.17583/remie.11765

Pritchard, D. (2010). Cognitive ability and the extended cognition thesis. *Synthese, 175*(S1), 133–151. https://doi.org/10.1007/s11229-010-9738-y

Pritchard, D. (2016). Intellectual virtue, extended cognition, and the epistemology of education. In J. S. Baehr (Ed.), *Intellectual virtues and education: Essays in applied virtue epistemology* (pp. 113–128). Routledge.

Pritchard, D., English, A. R., & Ravenscroft, J. (2021). Extended cognition, assistive technology and education. *Synthese, 199*(3–4), 8355–8377. https://doi.org/10.1007/s11229-021-03166-9

Rivera Novoa, Á. (2020). Mente extendida y transhumanismo: ¿Qué tan humana es la mente de un cyborg? In O. M. Donato Rodríguez, D. M. Muñoz González, & Á. Rivera Novoa (Eds.), *Redefinir lo humano en la era técnica perspectivas filosóficas* (75–89). Universidad Libre

Rivera-Novoa, A. (2024). The extended mind thesis and the transhumanist ideal of cognitive enhancement. *Trilogía Ciencia Tecnología Sociedad, 16*(33), e3142. https://doi.org/10.22430/21457778.3142

Rowlands, M. (2010). *The new science of the mind: From extended mind to embodied phenomenology*. MIT Press

Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. *Educational Psychology Review, 32*(1), 77–112. https://doi.org/10.1007/s10648-019-09498-w

Shafiq, M., Sami, M. A., Bano, N., Bano, R., & Rashid, M. (2025). Artificial intelligence in physics education: Transforming learning from primary to university level. *Indus Journal of Social Sciences, 3*(1), 717–733. https://doi.org/10.59075/ijss.v3i1.807

Skulmowski, A. (2024). Placebo or assistant? Generative AI between externalization and anthropomorphization. *Educational Psychology Review, 36*(2), 1–18. https://doi.org/10.1007/s10648-024-09894-x

Smart, P. (2017). Extended cognition and the internet: A review of current issues and controversies. *Philosophy & Technology, 30*(3), 357–390. https://doi.org/10.1007/s13347-016-0250-2

Smith, E. M., Graham, D., Morgan, C., & MacLachlan, M. (2023). Artificial intelligence and assistive technology: Risks, rewards, challenges, and opportunities. *Assistive Technology, 35*(5), 375–377. https://doi.org/10.1080/10400435.2023.2259247

Song, Y., Wu, K., & Ding, J. (2024). Developing an immersive game-based learning platform with generative artificial intelligence and virtual reality technologies – "LearningverseVR." *Computers & Education: X Reality, 4*, 100069. https://doi.org/10.1016/j.cexr.2024.100069

Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M., & Olson, C. B. (2024). Comparing the quality of human and chatgpt feedback of students' writing. *Learning and Instruction, 91*, 101894. https://doi.org/10.1016/j.learninstruc.2024.101894
Sterelny, K. (2010). Minds: Extended or scaffolded? *Phenomenology and the Cognitive Sciences, 9*(4), 465–481. https://doi.org/10.1007/s11097-010-9174-y
Sun, Y., Sheng, D., Zhou, Z., & Wu, Y. (2024). AI hallucination: Towards a comprehensive classification of distorted information in artificial intelligence-generated content. *Humanities and Social Sciences Communications, 11*(1), 1–14. https://doi.org/10.1057/s41599-024-03811-x
Suraj Kumar, S., Chatterjee, K., Ramchandra Reddy, S., Praveen Kumar, R., Sachin, A., Abhinav Vardhan Reddy, T., spsampsps Sriya Reddy, M. (2024). GenAI: Transforming climate science education in the digital age. In K. Raza, N. Ahmad, spsampsps D. Singh (Eds.), *Generative AI: Current Trends and Applications* (1177, 355–372). Springer Nature Singapore. https://doi.org/10.1007/978-981-97-8460-8_16
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.
Tang, K. (2024). Informing research on generative artificial intelligence from a language and literacy perspective: A meta-synthesis of studies in science education. *Science Education, 108*(5), 1329–1355. https://doi.org/10.1002/sce.21875
Tang, K.-S., & Cooper, G. (2024). The role of materiality in an era of generative artificial intelligence. *Science & Education*. https://doi.org/10.1007/s11191-024-00508-0
Tlili, A., Altinay, F., Huang, R., Bozkurt, A., Burgos, D., Shehata, B., Altinay, Z., Wang, H., spsampsps Sani-Bozkurt, S. (2024). Trends of artificial intelligence in special education and their relation to the sustainable development goals: Where are we now and where are we heading? In D. Burgos, J. W. Branch, A. Tlili, R. Huang, M. Jemni, C. M. Stracke, C. de la Higuera, C.-K. Looi, spsampsps K. Berrada (Eds.), *Radical solutions for artificial intelligence and digital transformation in education: Utilising disruptive technology for a better society* (27–45). Springer Nature. https://doi.org/10.1007/978-981-97-8638-1_3
UNESCO. (2023). *Guidance for generative AI in education and research*. https://unesdoc.unesco.org/ark:/48223/pf0000386693
Valverde-Berrocoso, J., Fernández-Sánchez, M. R., Revuelta Dominguez, F. I., & Sosa-Díaz, M. J. (2021). The educational integration of digital technologies preCovid-19: Lessons for teacher education. *PLoS ONE, 16*(8), 1–22. https://doi.org/10.1371/journal.pone.0256283
Vanhaelen, Q., Lin, Y.-C., & Zhavoronkov, A. (2020). The advent of generative chemistry. *ACS Medicinal Chemistry Letters, 11*(8), 1496–1505. https://doi.org/10.1021/acsmedchemlett.0c00088
Wang, K. D., Wu, Z., Tufts, L., Wieman, C., Salehi, S., & Haber, N. (2024). Scaffold or crutch? Examining college students' use and views of generative AI tools for STEM education (Version 1). *arXiv*. https://doi.org/10.48550/ARXIV.2412.02653
Wells, S. (2024). Ready or not, AI is coming to science education—And students have opinions. *Nature, 628*(8007), 459–461. https://doi.org/10.1038/d41586-024-01002-x
Wheeler, M. (2010). In defense of extended functionalism. In R. Menary (Ed.), *The extended mind* (pp. 245–270). The MIT press.
Wongvorachan, T., Lai, K. W., Bulut, O., Tsai, Y. S., & Chen, G. (2022). Artificial intelligence: Transforming the future of feedback in education. *Journal of Applied Testing Technology, 23*(1), 95–116.
Yang, B., He, L., Liu, K., & Yan, Z. (2024). VIAssist: Adapting multi-modal large language models for users with visual impairments. *arXiv*. https://doi.org/10.48550/arXiv.2404.02508
Ye, J. H., Zhang, M., Nong, W., Wang, L., & Yang, X. (2024). The relationship between inert thinking and ChatGPT dependence: An I-PACE model perspective. *Education and Information Technologies*, 1–25. https://doi.org/10.1007/s10639-024-12966-8
Zdravkova, K. (2022). The potential of artificial intelligence for assistive technology in education. In M. Ivanović, A. Klašnja-Milićević, spsampsps L. C. Jain (Eds.), *Handbook on intelligent techniques in the educational process* (29, 61–85). Springer International Publishing. https://doi.org/10.1007/978-3-031-04662-9_4
Zhai, X. (2022). ChatGPT user experience: Implications for education. *Available at SSRN 4312418*. https://doi.org/10.2139/ssrn.4312418
Zhang, Y., Sun, J., Feng, L., Yao, C., Fan, M., Zhang, L., Wang, Q., Geng, X., & Rui, Y. (2024). See widely, think wisely: Toward designing a generative multi-agent system to burst filter bubbles. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–24. https://doi.org/10.1145/3613904.3642545
Zielesny, A., Jain, L. C., & Kacprzyk, J. (2011). *From curve fitting to machine learning* (Vol. 18). Springer. https://doi.org/10.1007/978-3-642-21280-2