Obutchénie
Revista de Didática e Psicologia Pedagógica

DOI: http://doi.org/10.14393/OBv9.e2025-20

Artificial Intelligence in Education and the formation of consciousness: a Leontievian analysis of the risk of "personal sense" delegated to the algorithm[1]

A Inteligência Artificial na Educação e a formação da consciência: uma análise Leontieviana sobre o risco do "sentido pessoal" delegado ao algoritmo

Elisiane Spencer Quevedo Goethel[2]
Débora Cristina Fonseca[3]

ABSTRACT
This theoretical-critical essay explores the implications of the increasing integration of Artificial Intelligence (AI) systems in education, from the perspective of A. N. Leontiev's Activity Theory. The study discusses the paradox that while AI promises to optimize knowledge acquisition, it can simultaneously empty the personal meaning of study, an element fundamental for the formation of consciousness. It is argued that delegating the organization of learning or crucial task stages to the algorithm can transform the student's activity from a purpose-driven pursuit into a sequence of operational actions focused on algorithmic validation. The consequences of this transformation for the subjectivity of students and educators, as well as for the pedagogical relationship, are analyzed. The article concludes with a call for the development of a critical pedagogy and AI literacy, aiming to build a truly humanizing education, in which technology serves the formation of consciousness and not the inverse.

Keywords: Activity Theory. Artificial Intelligence in Education. Personal meaning. Critical AI Literacy.

RESUMO
Este artigo é um ensaio teórico-crítico que explora as implicações da crescente inserção de sistemas de Inteligência Artificial (IA) na educação, sob a ótica da Teoria da Atividade de A. N. Leontiev. O estudo discute o paradoxo de que, embora a IA prometa otimizar a aquisição de conhecimento, ela pode simultaneamente esvaziar o sentido pessoal do estudo, elemento fundamental para a formação da consciência. Argumenta-se que a delegação da organização da aprendizagem ou de etapas cruciais das tarefas ao algoritmo pode transformar a atividade do estudante de uma busca motivada por propósito em uma sequência de ações operacionais focadas na validação algorítmica. Analisam-se as consequências dessa transformação para a subjetividade de estudantes e educadores, bem como para a relação pedagógica. O artigo conclui com um apelo pela construção de uma pedagogia e um letramento críticos em IA, visando à construção de uma educação verdadeiramente humanizadora, na qual a tecnologia sirva à formação da consciência e não o inverso.

Palavras-chave: Teoria da Atividade. Inteligência Artificial na Educação. Sentido Pessoal. Letramento Crítico em IA.

[1] English version by Elisiane Spencer Quevedo Goethel. E-mail: lisispencer@gmail.com.

[2] Master in Education from the São Paulo State University "Júlio de Mesquita Filho" (Unesp) - Rio Claro Campus, São Paulo, Brazil. Orcid: https://orcid.org/0000-0002-6282-1563. E-mail: lisispencer@gmail.com.

[3] PhD in Psychology, Professor at the Department of Education/IB/Unesp Rio Claro. Orcid: https://orcid.org/0000-0001-8427-5973. E-mail: debora.fonseca@unesp.br.

DOI: http://doi.org/10.14393/OBv9.e2025-20

1 Introduction

The rise of artificial intelligence (AI) systems is one of the most significant technological milestones of our era. This ascent has profound implications across various sectors, including education. More than a simple tool or a set of algorithms, AI has become a constant presence that permeates and reconfigures social, economic, and, consequently, educational dynamics. The promise of personalization, optimization, and efficiency—for example, through the use of intelligent tutors that adapt content, recommendation systems for study paths, and platforms for instant performance feedback—has been a discourse frequently employed by publications not always specialized in the educational field, which tend to reinforce ideas of opportunity and advancement (LOURENÇO, 2023; LITWAK & GAMA, 2024; LIMA-LOPES, 2025). Digital platforms have become widespread, organizing learning trajectories based on algorithms that collect, analyze, and operate educational data in real-time. However, beyond the infrastructure of platforms and school management systems (POELL, NIEBORG & VAN DIJCK, 2019; LIMA-LOPES, 2025), there is an increasing use of AI as a direct tool by students to complete school activities and tasks, including text production and information retrieval (KOSMYNA et al., 2025; JELSON et al., 2025). The growing integration of AI into daily life, where it gradually becomes a primary executor of tasks, reinforces the need for a critical perspective (LIMA-LOPES, 2025).

This techno-scientific advancement, however, is not neutral. It carries with it a rationality that reconfigures the educational activity itself, potentially affecting the subjectivity of students and teachers and, ultimately, the very formation of consciousness. As emphasized by Leontiev (1978), human activity is the process that establishes the real lives of people, their social being in all its richness and variety of forms, and it is in this process that consciousness is formed (LEONTIEV, 1978). Activity Theory, developed by Leontiev, based on the foundations of dialectical historical materialism, offers a robust framework for analyzing this reconfiguration.

This article is structured as a theoretical-critical essay, grounded in a narrative review of the pertinent literature. Our objective is to problematize how algorithmic mediation can interfere with the formation of consciousness, in light

DOI: http://doi.org/10.14393/OBv9.e2025-20

of the risk of emptying personal meaning in the interaction with educational algorithms. To this end, the analysis is based on the articulation of cultural-historical psychology, focusing on the classical works of A.N. Leontiev on Activity Theory, and contemporary debates on the use of Artificial Intelligence (AI) in the educational context. The methodological approach involved the in-depth exploration of concepts from this perspective, as well as studies on the impacts of AI on cognitive and learning processes, and essays that problematize the ethical and pedagogical risks of its uncritical use. We therefore utilize contributions from authors such as Leontiev (1978), Vygotsky (2001), Davydov (1988), Saviani (2003), Frigotto (2010), Duarte (2001), and Libâneo (2012), among others. This approach aims to understand and propose pathways for the construction of a critical pedagogy and AI literacy, which preserve the subject's autonomy and the humanizing intentionality of education.

2 Foundations of Activity Theory and the Formation of Consciousness

Activity Theory proves relevant for understanding human interaction with technology, being fundamental for comprehending the relationship with Artificial Intelligence, as the relationship between human beings and tools has always been a central interest of this approach, which conceives it as a mediation process (KAPTELININ and NARDI, 2018). This premise was already present in Leontiev's works, which discussed various technologies, from the stone axe and the potter's wheel to the most advanced computational devices of his time.

Consciousness is not an innate given or an exclusively biological product, but rather a historical and social result of the subject's practical activity in the world (LEONTIEV, 1978). The maxim that man transforms reality while simultaneously transforming himself underscores the centrality of activity as a category of analysis. In this sense, activity is not merely a reaction to stimuli, but a complex and teleological process, driven by a motive that gives meaning to the action (LEONTIEV, 1978). As Eidt and Duarte (2007, p. 52) state, for cultural-historical psychology, "it is not the brain that creates life, but life that creates the brain", emphasizing the primacy of activity in the formation of the psyche.

DOI: http://doi.org/10.14393/OBv9.e2025-20

Leontiev (1978) distinguishes three hierarchical levels in the structure of human activity, each with its specific function and contribution to the formation of consciousness. The most comprehensive level and the basic unit of life is activity, impelled by a motive that satisfies a subject's need and gives it meaning. The motive of an action, often not directly conscious at its outset, is revealed by the emotional nuance that accompanies it and by an analysis of its dynamics and transformations over time. For Leontiev (1978), activity is the substance of consciousness, and he asserts that activity never occurs without a motive, even if it is subjectively or objectively hidden.

A second level is action, which corresponds to a process oriented towards a conscious goal that the subject seeks to achieve. Frequently, the purpose of an action differs from the motive of the activity that originates it. Leontiev (1978) illustrates this with the example of a person going to the library (action) to study a literary work (motive). The act of going to the library is guided by the objective of getting there, but the meaning of this action lies in the broader motive of knowledge appropriation. The author further emphasizes that activity and action are genuine realities but not coincident, as the same action can be performed in different activity contexts, demonstrating its relative independence.

Finally, the third level is operation, which refers to the ways or techniques by which an action is executed. Operations depend on the specific conditions in which the action is performed and frequently become automatized with practice, freeing consciousness to focus on other aspects of the activity (LEONTIEV, 1978). Leontiev exemplifies this distinction with the following example: to memorize a poem, the action is memorizing, but the operations (copying or mentally repeating) vary according to the conditions. The automatization of these operations, through practice and training, dispenses with the constant mediation of consciousness, transforming partial actions into a single, fluid action (LEONTIEV, 1978).

This macrostructure of human activity is dynamic and characterized by constant transformations. An activity can lose the motive that evoked it and become an action that performs a different activity, or, reciprocally, an action can acquire independent motivational force and become a special type of activity. It is

DOI: http://doi.org/10.14393/OBv9.e2025-20

important to remember that, according to Leontiev (1978), human activity is always directed by motives and mediated by instruments and signs, which are mediating artifacts. AI, in this context, emerges as a new and powerful mediating artifact. Its insertion into the classroom, whether as a platform or as a direct tool for the student, alters the conditions of learning activity, impacting students' actions and operations and, consequently, the very construction of meaning. An AI platform, for example, can offer instant feedback on a student's writing, leading them to revise their text and thus deepen their understanding of sentence structure. Similarly, an AI tool that suggests different approaches to solving a math problem can lead the student to explore new reasoning paths, altering how they operate in problem-solving.

The concept of personal meaning is central to understanding the formation of consciousness. For Leontiev, personal meaning "expresses the significance of reality and its parts for the subject himself" in terms of their history, needs, and social integration (LEONTIEV, 1978, p. 94). It differs from objective meaning, which is the socially established content of knowledge or phenomena. The appropriation of objective meanings, by itself, does not guarantee the attribution of personal meaning. Meaning emerges precisely from the subject's activity, their struggle to overcome difficulties, and the connection between the objective of their action and the intrinsic motive that drives them.

To illustrate this distinction and how personal meaning is constructed, let us consider an example from school everyday life: writing an essay on a complex topic, such as Climate Change. The apparent objective of the student's action is to produce a final text and achieve good grades. However, the implicit motive that truly gives meaning to this activity can be diverse, such as, for instance, the desire to express an opinion about what is happening to the environment, the pursuit of a deeper understanding of a global problem, or the satisfaction of mastering the art of writing and argumentation. If the student actively engages in research, critical analysis of sources, organization of their ideas, and formulation of their own arguments, overcoming the challenges of the writing process, the essay acquires a profound personal meaning. This active engagement transforms the

DOI: http://doi.org/10.14393/OBv9.e2025-20

task into an activity that drives the formation of their critical consciousness and intellectual autonomy (LEONTIEV, 1978; DUARTE, 2001).

On the other hand, when AI assumes and automates crucial stages of this activity—for example, generating a complete draft of an essay, summarizing complex information without requiring deep analysis, or offering ready-made corrections without demanding reflection on errors—it can prevent the student from experiencing the challenge and transformation that give rise to personal meaning. In this scenario, the student might achieve the external objective (submitting the essay), but the activity becomes a mere succession of operations without intrinsic purpose. The essence of appropriation, which is to reproduce the essential features of human activity accumulated in the object and in knowledge (LEONTIEV, 1978; EIDT and DUARTE, 2007), is weakened, as the student does not experience the process of construction.

Leontiev (1978) further argues that human needs are not given a priori but are developed by society and transformed through activity. He contends that productive activity, and not merely the satisfaction of needs, is the starting point and dominant element in human formation. The impoverishment of personality (LEONTIEV, 1978), can, under certain social conditions, combine with the satisfaction of a wide circle of everyday needs, representing a danger that consumer society poses to the human personality. This is exacerbated when AI, by delegating essential stages of activity, transforms the subject into a mere consumer of information and results, instead of an active producer of knowledge and meaning.

The formation of consciousness, therefore, is not a spontaneous process but the result of systematic processes of appropriation of historically constructed knowledge. Good teaching, as posited by Vygotsky (1991; 2001), is not just any teaching, but one that precedes and propels development, being a source of development and the emergence of the new (VIGOTSKI, 2001). The quality of psychological development thus depends on how teaching activity is organized to mobilize active appropriation, which leads to the emergence of new aptitudes and psychic functions (EIDT and DUARTE, 2007). Study activity, in particular, is

DOI: http://doi.org/10.14393/OBv9.e2025-20

considered the dominant activity in the developmental stage corresponding to school age, fundamental for the appropriation of the rudiments of the more developed forms of social consciousness, such as science, philosophy, and art (EIDT and DUARTE, 2007). The aptitude for logical thinking, for example, can only be the result of the appropriation of logic, a product of humanity's social practice (LEONTIEV, 1978).

It is important that artificial intelligence be employed as a tool for support and enhancement, and not as a substitute for the human role in education. When using AI, it is necessary to enrich the learning experience and deepen its meaning, going beyond the mere automation of tasks. In this way, it will become a powerful ally in building a more dynamic and engaging educational environment, where the teacher acts as a fundamental mediator in the process of discovery and construction of knowledge.

Therefore, the understanding and application of the fundamentals of Activity Theory are fundamental. Should the pedagogical process not preserve student protagonism, engagement in overcoming challenges, and the personal construction of knowledge, the use of AI risks diluting the personal meaning of study activity. This can lead to an impoverishment of human activity, transforming the student into a mere executor of tasks and distancing them from the true formation of consciousness. From this perspective, Artificial Intelligence, understood as a powerful cultural and mediating artifact, is not limited to a potential for emptying meaning. On the contrary, when consciously integrated into the pedagogical project, it can act as a valuable tool to expand the possibilities of human activity and, consequently, enrich the learning experience. AI can, for example, support the exploration of vast volumes of information or the simulation of complex scenarios, freeing the student to concentrate on critical analysis, formulating insightful questions, and constructing original syntheses. Thus, personal meaning is not delegated to the algorithm but emerges from overcoming amplified cognitive challenges mediated by technology, strengthening the mastery of actions and the formation of a subject's second nature, according to the principles of cultural-historical psychology. Criticism, therefore, is directed not at the technology itself, but at its uncritical use and the passive delegation of essential processes of human formation.

DOI: http://doi.org/10.14393/OBv9.e2025-20

For AI to truly contribute to the formation of active and transformative consciences, critical AI literacy becomes indispensable, empowering students to use technology ethically and strategically. A critically literate student can, for example, employ AI to generate different perspectives on a complex topic, without these replacing their own analysis, or to create initial drafts that will serve as a basis for their autonomous and original elaboration, maintaining the perception of authorship and deep engagement. Likewise, instead of accepting automated feedback as absolute truth, the critical student can use it as a starting point for reflection and the improvement of their ideas, fostering metacognition. Such approaches demonstrate that, instead of mere passive users, students can become active agents who shape AI to achieve meaningful learning objectives, transforming it into a powerful magnifying glass for critical thinking, rather than a crutch for cognitive laziness.

3 Artificial Intelligence and Subjectivity: Algorithmic Personalization and Direct Use as a Form of Mediation

The integration of artificial intelligence into educational processes is frequently associated with promises of personalization and efficiency, arguing that AI can adapt content and learning pace to each student's individual needs (LITWAK and GAMA, 2024; LOURENÇO, 2023). In this context, AI promises to revolutionize the learning experience through intelligent tutors and adaptive systems capable of offering individualized tutoring, adjusting content, pace, and challenges in real-time, or recommending personalized study paths. In terms of efficiency, the technology aims to provide immediate feedback on task performance, grammar, or style, in addition to automatically correcting exercises and assessments, speeding up the learning cycle and alleviating the teaching workload. Other frequently highlighted functionalities include the ability to generate text drafts, summarize complex content, or assist in research quickly and effectively, as well as integrating gamification elements to increase engagement and motivation.

DOI: http://doi.org/10.14393/OBv9.e2025-20

However, a critical analysis reveals that these same promises can, paradoxically, undermine the foundations of consciousness formation and the personal meaning of learning. The supposed algorithmic personalization, manifested in intelligent tutors, adaptive systems, or platforms that adjust content and pace, by determining ideal paths and adapting learning according to predictive patterns, risks limiting student autonomy in choosing motives and objectives of the activity. This dynamic can prevent exploratory seeking and the overcoming of challenges that give rise to personal meaning, transforming problem-solving into a sequence of reactive operations focused on algorithmic validation and replacing the intrinsic motive of mastering knowledge with the mere search for correct answers, which can lead to superficial learning. Similarly, AI is frequently used to assist in generating didactic content, such as creating personalized quizzes, article summaries, or presentation outlines. Although it promises to optimize time and efficiency, delegating the creation of these materials prevents student engagement in the critical synthesis of information and the formulation of their own questions, essential skills for knowledge appropriation.

Similarly, the promised efficiency of instant feedback and automated correction, while optimizing time, can divert the student's focus from deep understanding and reflection on error – essential for the formation of consciousness and the mastery of actions – to a simple search for validation and superficial correction. This compromises cognitive development and prevents deep engagement with learning material and the experience of knowledge construction. The promise of efficiency in generating text drafts, summaries, or research assistance, by delegating crucial stages of activity, can weaken memory, decrease the perception of authorship, and compromise active engagement in the critical synthesis of information and the formulation of one's own arguments. This functionality leads to a transition from active critical reasoning to passive supervision of generated content, potentially resulting in detriment to essential skills and the emptying of the personal meaning of the activity. Furthermore, gamification, by focusing on rewards and scores, can shift intrinsic learning

DOI: http://doi.org/10.14393/OBv9.e2025-20

motives towards superficial engagement driven by external stimuli, transforming learning into a process of consumption and performance for the algorithm, distancing it from the holistic formation of the subject.

It is necessary to critically question what is meant by personalization when decisions are made by algorithms, and what this efficiency signifies when AI tools perform tasks for the student (LIMA-LOPES, 2025). Recent studies have provided evidence of how the use of AI can alter cognitive processes. Research monitoring brain activity, for instance, demonstrates that neural connectivity in the brain systematically decreases with increased external support, such as that provided by Large Language Models[4] (LLMs). One study revealed that, in essay writing tasks, the group that used an LLM showed weaker neural connectivity and less engagement of alpha and beta networks in brain activities, in contrast to the group that performed the task using only their brain, without digital tool assistance, which exhibited stronger and more comprehensive networks (KOSMYNA et al., 2025). This is clear evidence of the transfer of part of the cognitive load to the automated system, which can result in long-term cognitive development compromise (KOSMYNA et al., 2025).

An inversion then occurs: instead of the student being the subject of learning who constructs their knowledge, they become an object of automated processes of content recommendation or commands to generate AI responses, where the student's behavioral past or the machine's processing capacity serves as the basis for future decisions. As Frigotto (2010) warns, this logic demonstrates the infiltration of instrumental rationality into the educational field, converting formation into training and knowledge into a commodity. This cognitive delegation to AI can lead to cognitive debt, where dependence on external systems replaces the mental effort required for independent thought. In this way, the uncritical use of AI in the educational environment can reinforce the subject's alienation from their study activity. Recalling Leontiev (1978), this is configured when actions cease to be internally motivated and

[4] Large Language Models (LLMs) are deep learning algorithms, trained with a large volume of data to understand and generate human language. ChatGPT is an example of an LLM.

DOI: http://doi.org/10.14393/OBv9.e2025-20

begin to respond to external commands, losing their connection to the activity as a whole. In other words, automated educational action does not form consciousness, but merely trains the response.

Libâneo (2012) observes that the centrality of the teacher as a mediator of learning is being eroded by technologies that promise to replace pedagogical mediation with self-executing teaching scripts or with tools that the student uses without proper critical guidance. This trend reduces the complexity of the teaching-learning relationship to simplified stimulus-response processes, disregarding the role of language, culture, and social contradictions in the formative process. Furthermore, algorithmic personalization and the uncritical use of AI promote curriculum fragmentation, with individual paths that disregard the collective, dialogue, and the historicity of knowledge. Education ceases to be a shared social practice and becomes a solitary experience, guided by performance goals. For Santaella (2019), this process compromises holistic formation, as it suppresses the aesthetic, ethical, and political dimensions of subjectivity. Educational subjectivity cannot be considered outside power relations and sociocultural mediations. By shifting the center of mediation to systems controlled by market interests, there is a risk of institutionalizing a new form of colonization of subjectivity, algorithmic colonialism (LIMA-LOPES, 2025). Thus, caution is needed when using AI in education. Algorithmic personalization and impulsive, uncritical use can depersonalize learning, replacing subjective experience with supposed technical neutrality. The challenge, therefore, is to place the subject at the center of the relationship with knowledge and to use technology as a means and not an end in education.

AI, by reducing cognitive load in tasks such as writing, can facilitate the comprehension and retrieval of information. However, this cognitive ease comes at a cost (STADLER et al, 2024; KOSMYNA et al., 2025). While it reduces mental effort, it can compromise germane cognitive load, which is the mental effort dedicated to constructing and automating robust learning schemas (STADLER et al, 2024; KOSMYNA et al., 2025). The reduction of this

DOI: http://doi.org/10.14393/OBv9.e2025-20

load can lead to less deep engagement and a transition from active critical reasoning to passive supervision of generated content (LEE et al., 2025; KOSMYNA et al., 2025). This raises serious concerns about students' ability to develop long-term critical thinking and problem-solving skills if they become accustomed to excessive dependence on AI solutions.

4 The Transformation of Study Activity and the Emptying of Personal Meaning

Study activity, from the perspective of cultural-historical psychology, transcends the mere execution of tasks or exposure to content. It is a specific form of activity oriented towards the conscious appropriation of systematized knowledge, essential for the formation of higher cognitive capacities. Not every teaching promotes psychic development, but good teaching anticipates development and "is a source of development and the emergence of the new" (VIGOTSKI, 2001, p. 334).

When this activity is mediated by algorithms that direct the student according to predictive patterns, or when the student delegates crucial stages of task execution to AI, it tends to deviate from its original formative purpose and, consequently, empty the personal meaning of learning. According to Davydov (1988), study needs to be planned as an activity for concept formation, in which the student engages with the content in order to understand its internal logic and social applicability. Learning, from this perspective, is not reduced to repetition or getting answers right, but involves the capacity to act intellectually upon the objects of knowledge. However, AI frequently organizes learning as a chain of micro-actions that prioritize task resolution based on statistical patterns or is used by the student to obtain ready-made answers, shifting attention from meaning to performance. This transformation is visible, for example, in educational platforms based on gamification, automated feedback, and extrinsic stimuli, or when AI is employed to generate texts, summaries, or problem solutions without the student's cognitive involvement. Corroborating this perspective, the study by Kosmyna et al.

DOI: http://doi.org/10.14393/OBv9.e2025-20

(2025) demonstrated the effects of this delegation: when using LLMs for essay writing, 83% of participants in the AI-assisted group failed to provide a correct citation in the first session, and none managed to cite correctly, evidencing difficulty in remembering and citing their own essays. The perception of essay ownership was also significantly lower for the LLM group compared to the "Brain-only" group. These results suggest that knowledge was not deeply internalized, evidencing superficial learning and a weakening of memory.

In this scenario of delegation, it is observed that, instead of acting as a mediator of thought, AI often assumes the role of executor, directly affecting the operations and actions of study activity. In programming classes, for example, students might resort to AI assistants to generate code snippets, debug errors, or find solutions to coding problems. While the tool can accelerate the process, uncritical use can prevent the student from experiencing the process of trial and error, logical reasoning, and problem debugging—fundamental stages for the development of logical thinking and problem-solving aptitude. Active engagement in intellectual production is replaced by passive supervision of generated content.

The student, accustomed to this dynamic, learns to act to receive rewards or avoid punishments, but does not necessarily understand the conceptual foundations of what they are doing. As Duarte (2001) points out, this practice demobilizes internal motivation and compromises the development of critical consciousness, as it transforms knowledge into a mere functional instrument. By reducing the mental effort to retrieve and curate information, AI can lead to a focus on verifying and modifying generated answers, instead of active engagement in intellectual production (LEE et al., 2025; KOSMYNA et al., 2025).

In Leontiev's conception (1978), activity imbued with personal meaning is one in which the subject recognizes their action as significant in their life trajectory, in their needs and projects. When this dimension is replaced by a system of automatic instructions or by the delegation of tasks to AI, the cultural mediation necessary for the constitution of subjectivity is lost. The action becomes an operation, and the subject, an executor.

DOI: http://doi.org/10.14393/OBv9.e2025-20

Leontiev (1978) demonstrates how the attribution of meaning transforms a seemingly trivial operation into a central activity for the personality. In the context of education, a powerful example emerges when a student engages in a project to make the school more accessible for a classmate with reduced mobility. The objective task might be to prepare a report or a plan for improvements, but the personal meaning for the student transcends academic obligation. It emerges from a deeper motive: empathy for the colleague, the genuine desire to promote inclusion, the satisfaction of applying knowledge to solve a real problem, and the perception of their ability to positively impact the environment and people's lives. This meaning is forged in the challenge to map architectural barriers, research accessibility solutions, dialogue with school administration, estimate costs, and, at times, deal with bureaucracy or indifference. It is in this active engagement, which involves critical reflection, collaboration, and overcoming obstacles, that their social consciousness and transformative autonomy are truly propelled.

In contrast, if AI is used to excessively automate and simplify crucial stages of this activity—for example, by generating a ready-made report with generic accessibility solutions, simulating barrier mapping without requiring careful observation and interaction with the real environment, or producing action plans without the need for the student to negotiate with different stakeholders—it can prevent the student from experiencing the challenge and transformation that give rise to personal meaning. In this scenario, the student might fulfill the external objective of submitting the plan, but the activity becomes a mere succession of fragmented operations.

Neurophysiological evidence corroborates this loss of meaning and engagement: the study by Kosmyna et al. (2025) revealed that the group of participants who used LLM and were then reallocated to writing without tools (LLM-to-Brain) showed weaker neural connectivity and less engagement of brain networks associated with planning and critical reasoning compared to those who always wrote without AI assistance (Brain-only). In the same study, it was observed that brain activity systematically decreased with the amount of external support.

DOI: http://doi.org/10.14393/OBv9.e2025-20

Additionally, the uncritical use of AI can accentuate the weakening of the teacher's role as an organizer of study activity, a trend already observed by Libâneo (2012). This reduction of teaching to mere supervision of tasks elaborated by technological devices, or to the inability to mediate the direct interaction of the student with AI, directly compromises dialogical mediation, an essential condition for the formation of theoretical thought and critical consciousness. The most serious consequence of this process is the emptying of human formation. As Saviani (2005) warns us, education needs to promote the overcoming of alienation, providing the subject with conditions to understand the world and intervene in it consciously. By transforming study into an activity regulated by algorithms and performance metrics, or delegated to the machine, AI in education can compromise this mission, naturalizing adaptation instead of transformation. Therefore, delegating the organization of learning or the execution of parts of the task to the algorithm represents not just a change in method, but a reconfiguration of the very logic of educational activity. This reconfiguration can threaten the constitution of personal meaning, the core of consciousness formation and the subject's autonomy.

5 Critical AI Pedagogy and Critical AI Literacy: Repositioning Technology in the Formation of Consciousness

The critique of algorithmic logic in education, as developed in this article, does not imply a simplistic or Luddite rejection of technology. On the contrary, it demands the development of a critical AI pedagogy that understands technology as a social and cultural mediation, intrinsically embedded in historical contradictions and power relations (LIMA-LOPES, 2025). In this sense, it is not merely about adapting teaching to AI, but about re-signifying the use of AI in light of the formative objectives of education, which should aim at human emancipation and the formation of active and transformative consciousnesses.

To this end, it is necessary to reassert the central role of the school as a space for intentional mediation between systematized knowledge and the subjects' experiences (SAVIANI, 2003). This mediation, which involves dialogue, reflection,

DOI: http://doi.org/10.14393/OBv9.e2025-20

and human interaction, cannot be entirely replaced by automatic systems that disregard the social, affective, and political links of learning. AI, therefore, should be seen as a cultural artifact (VIGOTSKI, 2001) that can expand the possibilities of mediation, provided it is under human pedagogical control.

Critical AI pedagogy must, thus, articulate three interconnected dimensions: the technical, the ethical, and the political. On the technical level, it implies understanding the functioning of AI platforms and tools, questioning algorithmic criteria, and contesting their parameters based on formative ends (LIMA-LOPES, 2025). On the ethical level, it demands respect for the autonomy of subjects, the plurality of knowledge, and data protection. On the political level, it requires a commitment to the formation of critical consciousness and the construction of a more just and democratic society.

Paulo Freire (1996) already warned us about the risks of banking education, where the student is a mere recipient of content. AI, if not carefully situated in the pedagogical project and critically mediated, can reinforce this logic, transforming the student into a consumer of predefined routes or a passive user of ready-made answers. In opposition, it is necessary to promote a problem-posing education, where the use of technology is subordinated to formative intentionality and dialogue with concrete reality. This implies recovering the value of the pedagogical act as an encounter between subjects, a mediation between knowledge, and the production of meanings. The algorithm can be a useful tool in this process, provided it does not replace the human bond that sustains the educational process. As Nóvoa (2021) states, pedagogy needs to reclaim its place in the face of technologies, not as a competitor, but as a critical horizon.

Fundamental to this pedagogy is critical AI literacy. This is not limited to the operational mastery of tools but encompasses the ability to understand, critically analyze, and use AI ethically and responsibly. For AI to be a tool that enriches and does not empty the meaning of learning, the student needs to be able to understand the nature of AI, evaluate it critically, and use it ethically and responsibly.

Understanding the nature of AI means recognizing that it is a constructed artifact that operates with algorithms and data, which may

DOI: http://doi.org/10.14393/OBv9.e2025-20

contain biases, and that it does not possess human consciousness, intentionality, or emotions (LIMA-LOPES, 2025).

Critically evaluating AI means questioning the information generated, identifying possible biases or distortions, verifying the veracity of data, and discerning between what is machine-generated and what is a product of human thought. The study by Kosmyna et al. (2025) demonstrated that, although AI could generate well-structured texts with good grammar, human teachers could identify a lack of personal meaning and originality, evaluating uniqueness differently. The homogeneity in essays generated by LLMs and the low statistical deviation compared to other groups indicate a standardization that can compromise the diversity and originality of thought (KOSMYNA et al., 2025).

Using AI ethically and responsibly means knowing when it is an appropriate tool for a specific task, respecting authorship—both human and AI-generated—data privacy, and avoiding misinformation. The study by Kosmyna et al. (2025) points out that dependence on AI, where students delegate cognitive and metacognitive responsibilities, can ultimately harm self-regulation and deep engagement with learning material.

The goal should always be for the student to be an active agent, not just a user, capable of shaping the use of AI to achieve meaningful learning objectives, recognizing that the intelligence and ultimate responsibility belong to the subject. Kosmyna et al. (2025) observed that while students with lower self-confidence tend to rely more on AI, students with higher competence use it strategically as an active learning tool (YANG et al., 2024). This strategic use is the manifestation of critical literacy in action, where AI is a powerful magnifying glass and not a crutch.

Critical AI literacy aligns with the need to educate individuals who act with freedom and autonomy (SAVIANI, 2003) and who fully develop their capacities.

A critical AI pedagogy also presupposes teacher training aimed at a critical reading of technologies. It is fundamental that educators are prepared to understand their foundations, limits, and potentials so that they can consciously intervene in their use.

DOI: http://doi.org/10.14393/OBv9.e2025-20

The challenges for this training are multiple, given the rapid evolution of AI, the complexity of its mechanisms, and the risk that, without proper preparation, teachers may also be led to a logic of instrumental use or to the reduction of their role as central mediators of learning, transforming into mere supervisors of automated scripts.

To overcome this threat, teacher training must go beyond operational mastery of tools. It needs to empower educators to understand the nature and mechanisms of AI, recognizing it as a constructed artifact with biases and without human consciousness or intentionality, so they can question its algorithmic criteria and contest its parameters based on formative ends. This implies preparing teachers to act as critical mediators of technology, guiding students to use AI as a powerful magnifying glass for critical thinking and the autonomous construction of knowledge, and not as a crutch that empties the personal meaning of activity. As Pimenta and Lima (2019) advocate, teacher training must articulate theory and practice to strengthen the professional identity and intellectual autonomy of educators, ensuring that AI serves the formation of active and transformative consciousnesses, and not the inverse.

Thus, instead of promoting a replacement of human mediations by technical logics, critical AI pedagogy, allied with students' critical literacy, proposes a repositioning: technology as an ally in the construction of meanings and not as a substitute for consciousness. Such a perspective refocuses the fundamental question at the center of the educational process: for whom, for what, and with what implications do we use artificial intelligence in education?

7 Final Considerations

This article sought to critically discuss the impacts of artificial intelligence on the formation of consciousness in the educational process, in light of Leontiev's Activity Theory. It was argued that the increasing delegation of pedagogical decisions or stages of study activity to algorithms threatens the personal meaning of study activity, by converting learning into a succession of operational actions disconnected from internal motivation and formative intentionality.

DOI: http://doi.org/10.14393/OBv9.e2025-20

From the cultural-historical perspective, consciousness is formed through activities guided by motives, mediated by culture and language. The replacement of pedagogical mediation or the uncritical use of AI by systems, when not reflected upon, compromises this dynamic, promoting a depersonalization of the educational experience. The analysis developed suggests that criticism, therefore, does not oppose technology itself, but proposes its conscious, ethical, and political employment, in favor of human emancipation.

The defense of a critical AI pedagogy and critical AI literacy emerges as fundamental to empower students to use technology as a resource to deepen learning, develop critical thinking, and sustain intellectual autonomy. This perspective demands an articulated and responsible action from educators, researchers, and public policy makers.

In this context, there is a need for the training and capacitation of educators to act as critical mediators, capable of discerning between uses of AI that enhance human activity and those that empty it. Their role is essential to foster an interaction that enriches the learning experience and promotes personal meaning, instead of passively delegating such processes. Furthermore, it is of utmost importance to expand research on the impact of AI on students' creativity, autonomy, and problem-solving capacity. The objective is to understand the long-term effects of cognitive delegation and to identify emancipatory pedagogical methodologies that incorporate artificial intelligence. In turn, public policy makers must consider approaches that transcend mere technical regulation, developing guidelines that ensure AI as a cultural artifact under human pedagogical control, aiming to protect subjectivity from potential algorithmic colonialism and invest in an educational infrastructure that prioritizes integral formation over instrumental efficiency.

In a scenario of increasing digitalization, the relevance of an education founded on praxis, dialogue, and the collective construction of meanings is reaffirmed. Thus, through this approach, it will be possible to reposition technology as a tool for meaning-making, and not as its substitute, contributing to a truly humanizing education. Artificial intelligence must, therefore, serve the formation of consciousness, and not the inverse.

DOI: http://doi.org/10.14393/OBv9.e2025-20

La Inteligencia Artificial en la Educación y la formación de la conciencia: un análisis leontieviano sobre el riesgo del "sentido personal" delegado al algoritmo

RESUMEN
Este artículo es un ensayo teórico-crítico que explora las implicaciones de la creciente inserción de sistemas de Inteligencia Artificial (IA) en la educación, bajo la óptica de la Teoría de la Actividad de A. N. Leontiev. El estudio discute la paradoja de que, aunque la IA prometa optimizar la adquisición de conocimiento, puede simultáneamente vaciar el sentido personal del estudio, elemento fundamental para la formación de la conciencia. Se argumenta que la delegación de la organización del aprendizaje o de etapas cruciales de las tareas al algoritmo puede transformar la actividad del estudiante de una búsqueda motivada por un propósito en una secuencia de acciones operacionales enfocadas en la validación algorítmica. Se analizan las consecuencias de esta transformación para la subjetividad de estudiantes y educadores, así como para la relación pedagógica. El artículo concluye con un llamado a la construcción de una pedagogía y una alfabetización críticas en IA, con el objetivo de edificar una educación verdaderamente humanizadora, en la cual la tecnología sirva a la formación de la conciencia y no lo inverso.

Palabras clave: Teoría de la Actividad. Inteligencia Artificial en la Educación. Sentido Personal. Alfabetización Crítica en IA.

8 References

DAVIDOV, V. V. La enseñanza escolar y el desarrollo psíquico. Moscú: Editorial Progreso, 1988.

DUARTE, Newton. Vigotski e o "aprender a aprender": crítica às apropriações neoliberais e pós-modernas da teoria vigotskiana. Campinas: Autores Associados, 2001.

EIDT, Nadia Mara e DUARTE, Newton. Contribuições da teoria da atividade para o debate sobre a natureza da atividade de ensino escolar. Psicol. educ. [online]. 2007, n.24, pp.51-72. ISSN 1414-6975. Disponível em: https://pepsic.bvsalud.org/scielo.php?pid=S1414-69752007000100005&script=sci_abstract. Acesso em: 17 jun. 2025

FRIGOTTO, Gaudêncio. A produtividade da escola improdutiva. 7. ed. São Paulo: Cortez, 2010.

FREIRE, Paulo. Pedagogia da autonomia: saberes necessários à prática educativa. São Paulo: Paz e Terra, 1996.

JELSON, A. et al. An Empirical Study to Understand How Students Use ChatGPT for Writing Essays. *arXiv preprint arXiv:2501.10551*, 2025. Disponível em: https://arxiv.org/abs/2501.10551.

KAPTELININ, Victor; NARDI, Bonnie. Activity Theory as a Framework for Human-Technology Interaction Research. Mind, Culture, and Activity, [s. l.], v. 25, n. 1, p. 3-5, 2018. DOI: https://doi.org/10.1080/10749039.2017.1393089.

KOSMYNINA, Natalya et al. Your Brain on ChatGPT: Accumulation of Cognitive Debt When Using an AI Assistant for Essay Writing Task. 2023. DOI: 10.48550/arXiv.2306.08872. Disponível em: https://arxiv.org/abs/2306.08872. Acesso em: 25 jun. 2025.

LEE, H.-P. et al. The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In: Proceedings of the ACM CHI Conference on Human Factors in Computing Systems. ACM, 2025. DOI: https://doi.org/10.1145/3706598.3713778.

LEONTIEV, A. N. O desenvolvimento do psiquismo. Lisboa: Horizonte Universitário, 1978.

LEONTIEV, A.N. Actividad, Conciencia y Personalidad. Ediciones Ciencias del Hombre. Buenos Aires: Argentina, 1978. (Obra original publicada em 1975).

LIBÂNEO, José Carlos. Pedagogia e pedagogos, para quê? São Paulo: Cortez, 2012.

LIMA-LOPES, Rodrigo Esteves de. Letramento de Dados e Suas Possibilidades Para a Educação Científica Em Linguagem. Revista Leia Escola, [s. l.], v. 23, n. 1, p. 62-81, 2023. Disponível em: https://zenodo.org/record/8021649. Acesso em: 26 jun. 2023.

LIMA-LOPES, Rodrigo Esteves de. Por uma Revisão Crítica do Uso de Inteligência Artificial na Educação. [s. l.]: SciELO Preprints, 2025. Disponível em: https://doi.org/10.1590/SciELOPreprints.11450. Acesso em: 24 jun. 2025.

LITWAK, Priscila; GAMA, Madson. Inteligência artificial revoluciona a forma de aprender: colégios adotam ferramenta para personalizar estudos. O Globo: Rio; Bairros, Rio de Janeiro, 28 set. 2024. Disponível em: https://oglobo.globo.com/rio/bairros/noticia/2024/09/28/inteligencia-artificial-revoluciona-a-forma-de-aprender-colegios-adotam-ferramenta-para-personalizar-estudos.ghtml. Acesso em: 7 jun. 2025.

LOURENÇO, Christine. Educação: inteligência artificial pode otimizar rotina pedagógica de instituições de ensino | Exame. Exame: Educação, São Paulo, 31 ago. 2023. Disponível em: https://exame.com/bussola/educacao-inteligencia-artificial-pode-otimizar-rotina-pedagogica-de-instituicoes-de-ensino/. Acesso em: 7 jun. 2025.

NÓVOA, António. Os professores depois da pandemia. Educação & Sociedade, Campinas, v. 42, 2021. DOI: https://doi.org/10.1590/ES.249236.

PIMENTA, Selma G.; LIMA, Maria Socorro Lucena. Estágio e docência: a relação necessária entre teoria e prática. São Paulo: Cortez, 2019.

POELL, Thomas; NIEBORG, David; VAN DIJCK, José. Platformisation. Internet Policy Review, [s. l.], v. 8, n. 4, 2019. Disponível em: https://policyreview.info/node/1425. Acesso em: 10 jun. 2025.

SANTAELLA, Lucia. Navegar no ciberespaço: o perfil cognitivo do leitor imersivo. São Paulo: Paulus, 2019.

SAVIANI, Dermeval. Escola e democracia. 38. ed. Campinas: Autores Associados, 2005.

SAVIANI, Dermeval. Pedagogia histórico-crítica: primeiras aproximações. 11. ed. Campinas: Autores Associados, 2003.

STADLER, M.; BANNERT, M.; SAILER, M. Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. Computers in Human Behavior, [s. l.], v. 160, p. 108386, 2024. DOI: https://doi.org/10.1016/j.chb.2024.108386.

VIGOTSKI, Lev S. A formação social da mente. São Paulo: Martins Fontes, 2001.

VIGOTSKI, Lev S. Aprendizagem e desenvolvimento intelectual na idade escolar. In: LURIA, A. R.; LEONTIEV, A. N.; VYGOTSKY, L. S. (orgs.). Psicologia e Pedagogia: bases psicológicas da aprendizagem e do desenvolvimento. São Paulo: Moraes, 1991.

VIGOTSKI, Lev S. A construção do pensamento e da linguagem. São Paulo: Martins Fontes, 2001.

YANG, Y. et al. Can We Delegate Learning to Automation?: A Comparative Study of LLM Chatbots, Search Engines, and Books. 2024. DOI: https://doi.org/10.48550/arxiv.2410.01396.

Received in June 2025
Approved in September 2025