Digital-Pedagogy.eu Institute for Education

Journal of Digital Pedagogy

2026, Vol. 5, No. 1, pp. 37-51

ISSN 3008-2021

doi.org/10.61071/JDP.2665

AI-Mediated Learning and the Restructuring of Interpretive Cognition: A Developmental-Critical Model for Social Sciences and Humanities Education

Cristina-Georgiana Voicu

voicucristina2004@yahoo.fr

Titu Maiorescu Secondary School, Iași, Romania

https://orcid.org/0000-0001-9299-6551

Abstract: This study examines how generative artificial intelligence (AI) reshapes interpretive cognition in Social Sciences and Humanities (SSH) education. Moving beyond instrumental perspectives, the paper conceptualizes AI as an epistemic mediator that reconfigures meaning-making processes, authorship, and learner agency. The study adopts a qualitative, conceptually driven design combining systematic interdisciplinary synthesis with structured classroom observation (N = 42 interactional episodes, 27 students, 6 instructional sessions). An abductive analytical approach was used to identify patterns of learner – AI interaction and to construct a developmental-critical model of interpretive learning. Findings reveal three interconnected transformations: the externalization of interpretive cognition, the hybridization of authorship, and the emergence of distributed epistemic agency. These correspond to three developmental trajectories – AI-dependent, AI-enhanced, and AI-critical interpretation – reflecting increasing levels of interpretive autonomy. The study argues that AI does not merely support learning but fundamentally restructures the cognitive and epistemic architecture of interpretation. In response, a developmental-critical pedagogical model is proposed, operationalized through six checkpoints designed to scaffold critical AI literacy, epistemic responsibility, and interpretive autonomy. The article contributes to ongoing debates on AI in education by providing a theoretically grounded and empirically informed framework for integrating AI into SSH pedagogy without compromising interpretive depth.

Keywords: AI-mediated learning; interpretive cognition; epistemic agency; critical AI literacy; digital pedagogy; SSH education; generative AI

Received: 14.04.2026. Accepted and published: 05.06.2026

Citation:

Voicu, C.-G. (2026). AI-Mediated Learning and the Restructuring of Interpretive Cognition: A Developmental-Critical Model for Social Sciences and Humanities Education. *Journal of Digital Pedagogy, 5*(1) 37-51. Bucharest: Institute for Education. https://doi.org/10.61071/JDP.2665

1. Introduction

Artificial intelligence has evolved from an infrastructural technology into an active cognitive agent within educational environments. Generative AI systems now perform tasks central to SSH education, including textual interpretation, argument construction, and contextual explanation (Zhai et al., 2024; Wang et al., 2024). Processes once grounded in sustained cognitive effort and dialogic reasoning are increasingly compressed into instantaneous outputs, transforming the conditions under which learners engage with meaning.

From a hermeneutic perspective, interpretation is not merely the extraction of meaning but a situated, historically embedded act shaped by the interpreter's horizon of understanding (Gadamer, 2004). AI-generated interpretations, however, operate outside such lived horizons, producing decontextualized yet rhetorically coherent outputs, thereby creating a tension between interpretation as lived understanding and interpretation as computational simulation (Floridi, 2023). This transformation raises fundamental pedagogical and epistemological questions concerning interpretive agency, authorship, and cognitive development (Selwyn, 2023; Giannakos et al., 2025), and calls for a developmental-critical framework integrating cognitive, ethical, and pedagogical dimensions of AI-mediated learning.

Figure 1

The Changing Landscape of Interpretation in AI-Mediated SSH Education

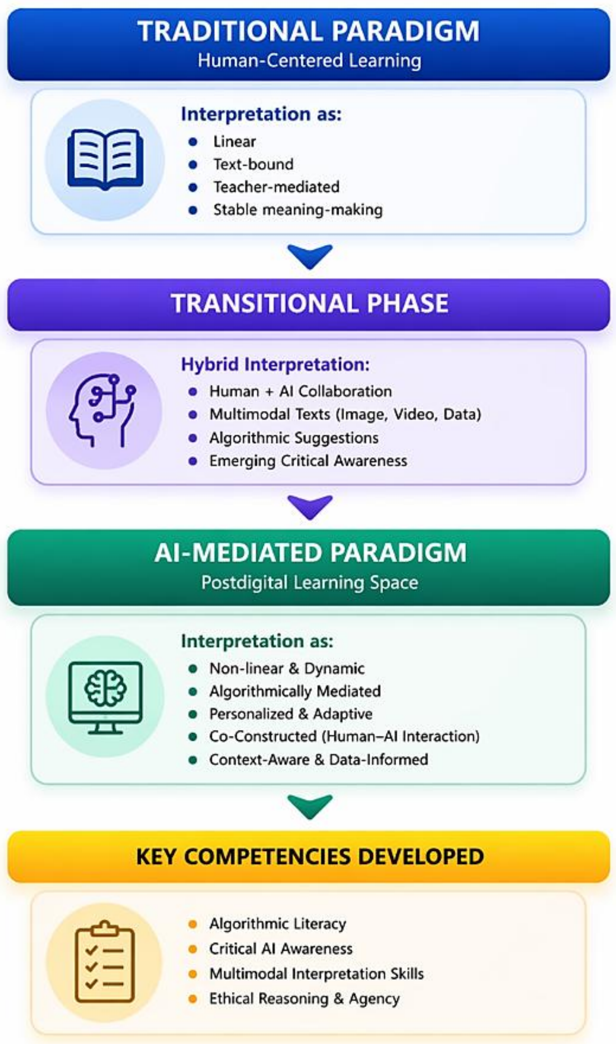


Figure 1 illustrates how interpretive cognition is reshaped within AI-mediated SSH education. At the centre lie reading, inference, and contextualization, while outer layers represent increasing AI involvement, from assisted sense-making and summarization to generative sense-making and cognitive acceleration. The broadest layer captures systemic consequences such as cognitive shortcuts, hybrid authorship, distributed epistemic agency, and iterative response in contemporary SSH learning ecologies

1.1. The Cognitive-Epistemic Shifts Introduced by AI

Artificial intelligence increasingly functions as an epistemic technology that generates outputs resembling interpretation while lacking genuine contextual awareness or lived experience. Although such systems produce coherent and analytically structured responses, they do not engage in interpretation as a situated, historically grounded act; instead, they simulate reasoning through probabilistic pattern recognition. This shift raises important questions about the quality and nature of meaning-making in AI-mediated environments.

Three interconnected shifts become evident. First, interpretive processes undergo cognitive acceleration, as AI compresses extended reasoning into instantaneous outputs. Second, AI introduces a simulation of interpretation, performing acts such as summarization, contextualization, and analysis without reflexivity or situated judgment. Third, hybrid authorship emerges as human and AI-generated reasoning become increasingly entangled in student work. Taken together, these developments reshape the trajectory of interpretive cognition and reinforce the need for critical pedagogical mediation, especially given evidence that generative AI is structurally transforming teaching and learning practices (Xiaoyu et al., 2025).

1.2. AI and Digital Learning Ecologies

The integration of artificial intelligence into education has generated complex digital learning ecologies in which cognition is distributed across networks of human and technological actors. Learning no longer unfolds solely within the individual mind but through continuous interaction with AI systems that provide scaffolding, feedback, and interpretive guidance.

While these affordances can enhance accessibility and engagement, they also create challenges related to epistemic authority and dependency. Foundational work anticipated these systemic transformations in education (Holmes et al., 2019; Luckin & Cukurova, 2019), and more recent reviews confirm their impact on pedagogical structures and epistemic habits (Monib et al., 2024). Within such ecologies, the teacher's role shifts from knowledge provider to facilitator of epistemic processes, guiding learners in interpreting AI outputs, questioning their validity, and integrating them critically into disciplinary frameworks.

1.3. A Developmental-Critical Perspective

Learning is conceptualized as a gradual process grounded in reflection, interaction, and contextual engagement rather than the rapid acquisition of ready-made outputs.

Cognitive-developmental theories emphasize the structured internalization of reasoning processes (Piaget, 1972; Vygotsky, 1978), while hermeneutic perspectives define interpretation as dialogic and situated (Gadamer, 2004; Ricoeur, 1981). These are complemented by distributed cognition theory (Hutchins, 1995) and post-digital perspectives highlighting the entanglement of human and technological actors (Jin et al., 2025). Together, these frameworks support the need for pedagogical models that preserve interpretive agency under AI mediation.

1.4. The Need for a New Model

Although research on AI in education has expanded rapidly, existing models often emphasize functional, ethical, or technological dimensions while overlooking the specific epistemic demands of SSH disciplines. In these contexts, learning cannot be reduced to efficiency or performance optimization, because it depends on sustained engagement with ambiguity, interpretation, and contextual reasoning.

This limitation is especially visible in higher education research, where generative AI is frequently framed through institutional adaptation and opportunity rather than interpretive development (Bobula, 2024). Systematic reviews also suggest that the literature remains largely focused on implementation and performance outcomes, leaving disciplinary specificity underexplored (Garzón, 2025). The present study responds to this gap by proposing a developmental-critical model organized around stages of interpretive development and pedagogical checkpoints designed to support critical AI engagement.

1.5. Research Questions

To guide the conceptual and analytical development of the study, the following research questions are proposed:

RQ1. How does AI-mediated learning reshape interpretive cognition in Social Sciences and Humanities (SSH) education?

RQ2. What types of epistemic agency emerge in AI-mediated learning environments, and how are they distributed between human and algorithmic actors?

RQ3. In what ways does the integration of generative AI influence authorship, argumentation, and meaning-making processes in SSH contexts?

RQ4. What developmental trajectories characterize learners' engagement with AI in interpretive tasks?

RQ5. How can pedagogical frameworks be designed to support the transition from AI-dependent to AI-critical interpretation?

2. Theoretical Framework

Understanding AI-mediated learning requires moving beyond purely technical or instrumental perspectives and engaging with the epistemological transformations it entails. In SSH disciplines, meaning-making has traditionally been grounded in hermeneutic frameworks that emphasize dialogue, context, and interpretive agency, where understanding emerges through the interaction between text, context, and subjectivity. Generative AI disrupts this dynamic by producing structurally coherent but fundamentally decontextualized outputs that simulate understanding without experiential grounding (Floridi, 2023).

From a cognitive-developmental perspective, learning is a gradual process of internalization supported by effort, reflection, and social interaction (Piaget, 1972; Vygotsky, 1978). AI-mediated environments create the risk of premature cognitive externalization, as learners may rely on algorithmic outputs before developing the internal structures necessary for independent reasoning. At the same time, distributed cognition theory highlights how knowledge production becomes extended across human and technological systems (Hutchins, 1995), creating hybrid forms of agency that can enhance cognition but also encourage epistemic outsourcing when learners defer uncritically to AI-generated outputs (Selwyn, 2023). These tensions converge in the need for critical AI literacy and motivate the developmental-critical model proposed in this study.

2.1. AI as an Epistemic Technology

Artificial intelligence can be understood as an epistemic technology that generates interpretations without possessing subjectivity or contextual awareness (Floridi, 2023). Unlike human interpreters, whose understanding is historically and experientially grounded, AI produces outputs through statistical inference, simulating reasoning without engaging in it as a meaningful cognitive act.

AI-generated outputs, although coherent, remain detached from these dimensions, creating an epistemic gap between simulated and situated understanding.

2.2. Meaning-Making in SSH

Meaning-making in SSH involves interpretive reasoning, argumentation, and ethical reflection, requiring sustained engagement with texts and contexts. AI disrupts these processes by providing immediate, structured outputs that may replace rather than support cognitive effort.

While such outputs can facilitate access to information, they risk bypassing the developmental processes through which understanding is constructed, requiring careful pedagogical mediation.

2.3. Cognitive Development Under AI Mediation

Cognitive-developmental theory conceptualizes learning as the gradual internalization of reasoning through interaction and reflection (Piaget, 1972; Vygotsky, 1978). AI-mediated environments introduce the risk of premature cognitive externalization, where learners rely on outputs before developing the underlying structures of reasoning.

This may lead to reduced metacognitive engagement and diminished interpretive depth (Karmiloff-Smith, 1992), highlighting the need for AI to function as scaffold rather than substitute.

2.4. Distributed Cognition and Algorithmic Co-Agency

Distributed cognition theory explains how cognitive processes extend beyond the individual to include tools and environments (Hutchins, 1995). In AI-mediated learning, this includes algorithmic systems that participate in meaning-making.

While this distribution can enhance learning, it also introduces risks of epistemic outsourcing when learners defer uncritically to AI-generated outputs (Selwyn, 2023), complicating responsibility and knowledge validation.

2.5. Critical AI Literacy

Critical AI literacy extends beyond technical competence to include epistemic and ethical awareness. It enables learners to evaluate the validity, limitations, and implications of AI-generated knowledge.

This literacy encompasses technical, epistemic, and ethical dimensions and is essential in SSH contexts where interpretation is central (Daud et al., 2025).

2.6. Post-digital Perspectives

Post-digital theory provides a lens for understanding AI as embedded in contemporary learning environments rather than as an external tool. Human and technological processes become deeply intertwined, making AI an integral component of the learning ecology rather than an optional addition.

This entanglement challenges traditional distinctions between human and technological agency, as learning increasingly emerges from interactions between learners and algorithmic systems. Post-digital perspectives therefore call for adaptive and reflexive pedagogical frameworks attentive to the evolving dynamics of AI-mediated environments.

2.7. Toward a Developmental-Critical Model

The theoretical perspectives outlined above converge in the need for a framework integrating cognitive development, epistemic agency, and pedagogical design. The developmental-critical model proposed in this study is grounded in three propositions: interpretation is inherently situated and human-centred, cognitive development depends on reflective engagement, and pedagogy must actively scaffold epistemic responsibility.

Within this framework, AI is understood not as a replacement for human reasoning but as a mediator that must be critically integrated into learning processes. The model seeks to guide learners toward increasing interpretive autonomy while preserving critical engagement.

3. Methods

This study adopts a qualitative, conceptually driven design grounded in a post-phenomenological perspective (Jin et al., 2025), conceptualizing AI as an epistemic mediator of interpretive processes. The empirical component consists of structured classroom observations conducted across six instructional sessions involving 27 lower secondary students (aged 13-15).

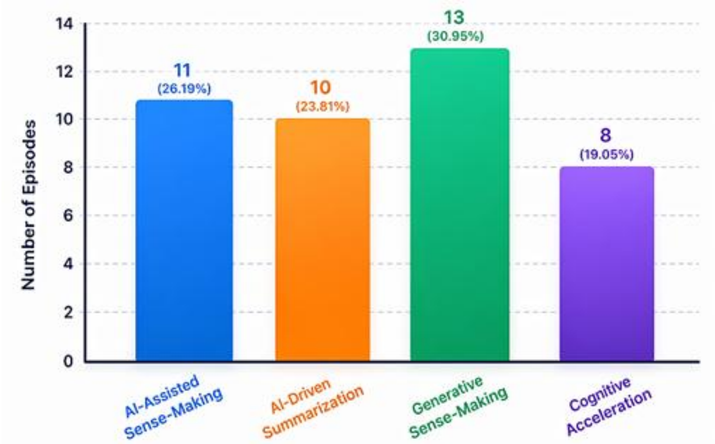
Table 1

Distribution of Interactional Episodes by Type of AI Involvement

Category of AI Involvement	Frequency (n)	Percentage (%)
AI-assisted sense-making	11	26.19%
AI-driven summarization	10	23.81%
Generative sense-making	13	30.95%
Cognitive acceleration	8	19.05%
Total	**42**	**100%**

Figure 2

Distribution of AI Involvement Across the 42 Observed Interactional Episodes



To ensure transparency of the empirical distribution, **Table 1** shows the frequency and percentage of interactional episodes across categories of AI involvement. The observational corpus includes **42 interactional episodes**, defined as discrete instances of learner-AI engagement during interpretive tasks (text analysis, argument construction, contextual explanation). Each episode was systematically coded based on: type of AI involvement, degree of cognitive delegation, epistemic positioning of the learner. Episodes were categorized into four analytical groups: 1. AI-assisted sense-making; 2. AI-driven summarization; 3. Generative sense-making; 4. Cognitive acceleration. The coding process followed an abductive approach, combining inductive pattern identification with theoretical alignment (Piaget, 1972; Hutchins, 1995). Frequencies were calculated across all episodes to identify dominant interaction patterns, which were subsequently interpreted in relation to expected categories derived from the literature. This enabled the construction of developmental trajectories reflecting increasing levels of interpretive autonomy.

Figure 2 shows the distribution of 42 coded interactional episodes across four categories of AI involvement. Generative sense-making is the most frequent category (30.95%), followed by AI-assisted sense-making (26.19%) and AI-driven summarization (23.81%), while cognitive acceleration is least frequent (19.05%). The distribution indicates that AI is predominantly used as a support for interpretive engagement rather than as a full substitute for cognitive processing. At the same time, the presence of summarization and cognitive acceleration episodes indicates that some learners relied on AI in ways that reduced direct interpretive effort. The observed pattern therefore points to a transitional learning ecology in which assisted interpretation is dominant, but risks of epistemic outsourcing are already visible. These distributions derive directly from the 42 coded interactional episodes described in Sections 3.3 and 3.5.

3.1. Post-Phenomenological Orientation

This study is grounded in a post-phenomenological perspective, which conceptualizes technologies as mediators of human experience rather than neutral tools (Floridi, 2023). Artificial intelligence is thus understood as shaping perception, interpretation, and action, influencing how learners engage with knowledge in educational contexts.

Post-phenomenology emphasizes the relational nature of technological mediation, avoiding deterministic interpretations of AI as inherently beneficial or detrimental. Instead, it highlights how its effects depend on interactions between users, contexts, and practices (Hutchins, 1995; Selwyn, 2023). In SSH education, where interpretation is central, this perspective is especially useful because AI does not replace interpretive processes but reconfigures how learners engage with texts, construct meaning, and position themselves epistemically.

3.2. Context and Data Sources

The empirical component of the study was conducted in a lower secondary educational setting in Iași, Romania (Titu Maiorescu Secondary School), with one class of 27 students aged 13-15. Observations were carried out across six instructional sessions over a six-week period, with one observed session per week, integrated into regular classroom activities.

The sessions focused on key meaning-making tasks specific to Social Sciences and Humanities (SSH) education, including textual interpretation, argumentative writing, contextual explanation, and ethical reflection. These task types were selected because they allow close examination of interpretive reasoning under AI mediation.

The observational corpus consists of 42 interactional episodes, defined as discrete sequences of learner-AI engagement occurring within clearly identifiable interpretive tasks. An episode was recorded when students used a generative AI system to support, replace, extend, or accelerate a meaning-making process, such as identifying themes, generating summaries, contextualizing ideas, or formulating arguments.

The study did not evaluate individual student performance; instead, it examined recurrent patterns of interaction, focusing on how learners positioned AI within their reasoning processes, how they incorporated AI-generated outputs, and how epistemic responsibility shifted during task completion.

3.3. Analytical Procedure

The analysis followed an abductive interpretive strategy, combining inductive observation of classroom interaction with theoretically informed categorization. The coding process unfolded in three stages. First, all 42 interactional episodes were reviewed and described in relation to task type, student use of AI, and the role of AI output in the completion of the task. Second, episodes were coded according to three analytical dimensions: type of AI involvement, degree of cognitive delegation, and epistemic positioning of the learner. Third, the coded patterns were interpreted through cognitive-developmental, hermeneutic, and distributed cognition frameworks and then consolidated into broader developmental trajectories.

Field notes were produced during and immediately after each session using a structured observation grid focused on task type, prompt use, AI function, learner elaboration, and epistemic stance. Coding was subsequently conducted across the full corpus in iterative cycles of comparison and category refinement.

To make the coding logic explicit, each episode was first recorded descriptively and then assigned to one dominant category of AI involvement according to its primary observable function in the task. Episodes were subsequently aligned with one of the three developmental trajectories on the basis of four interpretive indicators: degree of learner elaboration, presence or absence of verification, extent of reformulation, and evidence of critical engagement with the AI output. Episodes characterized by direct uptake with minimal elaboration were classified as AI-dependent; episodes involving selective adaptation or partial reformulation were classified as AI-enhanced; and episodes involving triangulation, questioning, contextual reframing, or explicit evaluation of the AI response were classified as AI-critical.

To increase analytical transparency, coding decisions were guided by repeated observable indicators such as whether AI was used to support interpretation, generate condensed content, provide meaning directly, or accelerate task completion with minimal learner elaboration. The resulting pattern distribution was then used as the empirical basis for the developmental-critical model proposed in this study. The resulting coded distributions form the empirical foundation for the quantitative summaries presented in Figures 2 and 3.

3.4. Component I: Systematic Conceptual Synthesis

The first component of the methodological design consists of a systematic conceptual synthesis of interdisciplinary literature. Rather than following a narrowly defined systematic review protocol, this approach integrates insights from multiple fields to construct a comprehensive analytical framework drawing on cognitive science, AI ethics, digital pedagogy, philosophy of interpretation, and developmental psychology.

Recent bibliometric analyses confirm the rapid consolidation of generative AI as a major research domain in higher education while highlighting the need for more conceptually differentiated pedagogical models (Dai, 2026; Întorsureanu et al., 2025). The synthesis is organized around key theoretical pillars, including distributed cognition (Hutchins, 1995), cognitive-developmental theory (Piaget, 1972; Vygotsky, 1978), AI literacy and ethics (Zhai et al., 2024; Selwyn, 2023), and hermeneutic traditions (Gadamer, 2004; Ricoeur, 1981).

3.5. Component II: Classroom Narrative Observation

The classroom observation component was conducted with 27 lower secondary students in Iași, Romania, across six instructional sessions embedded in regular SSH-related learning activities. The sessions included tasks involving textual interpretation, argumentative writing, contextual explanation, and ethical discussion. AI-supported interaction was observed whenever students consulted a generative AI system to clarify a text, obtain a summary, generate interpretive ideas, formulate arguments, or accelerate task completion. Across the six sessions, 42 interactional episodes were identified and retained for analysis.

The observation protocol focused on five concrete elements: the task being performed, the prompt or request addressed to AI, the function played by the AI response, the degree of learner elaboration after receiving the output, and the epistemic stance adopted toward the generated content. Episodes were categorized into four types of AI involvement: AI-assisted sense-making, when AI supported a student's own interpretive effort; AI-driven summarization, when AI condensed content or replaced direct engagement with the source material; generative sense-making, when AI actively produced interpretive or argumentative material used in the task; and cognitive acceleration, when AI was used primarily to shorten effort, speed up completion, or bypass intermediate reasoning steps. Categorization was based on the dominant observable function of AI in each episode.

No identifiable personal data were collected, and all observations were recorded at the level of classroom interaction rather than individual profiling.

For example, an episode was coded as AI-dependent when a learner requested an instant interpretation or summary, accepted the generated response with minimal modification, and used it directly in task completion without verification or contextual expansion. By contrast, an episode was coded as AI-critical when a learner compared the AI output with the source text, questioned omissions or inaccuracies, reformulated the response in their own words, or rejected parts of the generated content as insufficient for the task. These contrasts helped stabilize the interpretive distinctions used in the analytical model.

To further clarify the empirical basis of the study, the 42 episodes were distributed across the observed task types as follows: 14 episodes involved textual interpretation, 11 involved argumentative writing, 9 involved contextual explanation, and 8 involved ethical reasoning activities. This distribution indicates that the observational corpus captures AI-mediated interaction across multiple core SSH practices rather than a single isolated use case.

Ethical considerations. The study involved non-interventional classroom observation conducted during regular instructional activities. No identifiable personal data were collected, and all analytical notes were anonymized at the point of recording. The study focused on interactional patterns rather than individual profiling and was conducted in accordance with institutional ethical expectations for educational research.

Table 1a

Analytical Criteria Used to Relate AI Involvement Categories to Developmental Trajectories

AI involvement category	Primary observable function	Typical learner behaviour	Most likely developmental trajectory
AI-assisted sense-making	AI supports ongoing interpretation	Student uses AI to clarify, extend, or test emerging understanding	AI-enhanced
AI-driven summarization	AI condenses or substitutes source engagement	Student relies on shortened output instead of sustained reading	AI-dependent
Generative sense-making	AI produces interpretive or argumentative content	Student incorporates AI-generated ideas with varying degrees of reformulation	AI-dependent or AI-enhanced
Cognitive acceleration	AI shortens effort or bypasses intermediate reasoning	Student prioritizes speed and task completion over interpretive depth	Usually AI-dependent, unless critically regulated

This analytical crosswalk was used heuristically rather than mechanically, as final trajectory assignment depended on the observable degree of learner autonomy, verification, and interpretive control within each episode. The empirical distribution of the observed episodes is presented in Table 1 and Figure 2, while the analytical criteria used to relate categories of AI involvement to developmental trajectories are summarized in Table 1a.

3.6. Component III: Interpretive-Developmental Modelling

The third component involves the construction of an interpretive-developmental model through an abductive analytical process. Abduction makes it possible to bridge empirical observations and theoretical constructs, enabling the identification of meaningful patterns in AI-mediated learning.

The modelling process is guided by three analytical categories – interpretive cognition, epistemic agency, and critical AI literacy – and is aligned with established cognitive-developmental theories, including those of Piaget (1972), Vygotsky (1978), Karmiloff-Smith (1992), and Kuhn (1991). This ensures that the model is both empirically grounded and theoretically robust, reflecting recognized trajectories of cognitive and epistemic development.

3.7. Methodological Rigor

Ensuring methodological rigor is a central concern of this study, particularly given its qualitative and conceptual orientation. One key strategy is theoretical saturation, achieved through the integration of diverse interdisciplinary perspectives that collectively support a coherent analytical framework.

Ecological validity is also emphasized, as the study is grounded in real classroom contexts where AI-mediated learning unfolds naturally. Finally, reflexive analysis helps maintain conceptual clarity by avoiding the anthropomorphization of AI systems and ensuring that they are understood as technological mediators rather than autonomous agents (Floridi, 2023; Selwyn, 2023).

4. Results

The analysis reveals three interrelated macro-transformations characterizing AI-mediated SSH learning: the externalization of interpretive cognition, the hybridization of authorship and argumentation, and the emergence of distributed epistemic agency. These transformations arise from recurring interactional patterns and collectively reshape the developmental trajectory of learners.

The externalization of interpretive cognition is visible in students' increasing reliance on AI for summarization and contextual explanation, which enhances efficiency but reduces engagement with interpretive processes and may lead to shallow learning (Wang et al., 2024). Hybridization of authorship reflects the blending of human and AI-generated content, often producing standardized rhetorical structures and reduced originality (Selwyn, 2023; Floridi, 2023). Distributed epistemic agency further complicates this landscape, as learners attribute authority to AI outputs and shift responsibility for knowledge validation (Giannakos et al., 2025), giving rise to three trajectories: AI-dependent, AI-enhanced, and AI-critical interpretation.

4.1. Externalization of Interpretive Cognition

One of the most salient transformations in AI-mediated SSH learning is the externalization of interpretive cognition. Students increasingly delegate thematic identification, contextual explanation, and comparative analysis to AI systems, reflecting broader dynamics of cognitive acceleration and aligning with both dual-process theories of cognition (Kahneman, 2011) and research on cognitive offloading in digital environments (Wang et al., 2024).

Although such outputs may be coherent and structurally sound, they reduce opportunities to engage with ambiguity, complexity, and textual nuance. This also erodes the productive struggle essential for cognitive development, which hermeneutic and developmental perspectives associate with reflection and revision (Ricoeur, 1981; Karmiloff-Smith, 1992), and may encourage shallow reading practices and contextual drift (Carr, 2010).

4.2. Hybridization of Authorship and Argumentation

A second major transformation concerns the hybridization of authorship and argumentation, as students increasingly produce texts that blend human and AI-generated content. This blending often occurs without explicit awareness of the boundary between personal reasoning and algorithmic output, raising questions about intellectual ownership and academic integrity in AI-mediated environments (Selwyn, 2023).

One significant consequence is the normalization of argumentative patterns. AI-generated texts tend to follow statistically dominant rhetorical structures, producing coherent yet often formulaic arguments. As students rely on such outputs, writing may converge toward standardized forms, limiting diversity of expression and constraining the development of individual voice (Floridi, 2023). In addition, learners may incorporate AI-generated evidence without critically evaluating its accuracy, sometimes including fabricated or unverifiable references (Zhai et al., 2024).

4.3. Distributed Epistemic Agency

The third transformation identified in the analysis is the emergence of distributed epistemic agency within AI-mediated learning environments. Drawing on distributed cognition theory, knowledge production is no longer confined to the individual learner but is distributed across networks that include technological systems (Hutchins, 1995), with AI functioning as an epistemic participant that influences how knowledge is generated and validated.

Students' interactions with AI reveal a tendency to attribute authority to algorithmic outputs, treating them as reliable sources of knowledge. Expressions such as "the AI says" reflect a shift toward machine authority, where AI-generated responses are seen as credible, sometimes even surpassing traditional academic sources (Selwyn, 2023; Giannakos et al., 2025). This redistribution of authority is accompanied by a displacement of responsibility, as learners may relinquish the task of evaluation, leading to epistemic outsourcing (Zhai et al., 2024). At the same time, some learners engage dialogically with AI, using it as a tool for inquiry and reflection rather than as an unquestioned authority.

4.4. Developmental Trajectories

The coded interactional patterns were further consolidated into three developmental trajectories of AI-mediated interpretation: AI-dependent interpretation, AI-enhanced interpretation, and AI-critical interpretation. Of the 42 observed episodes, 15 were classified as AI-dependent, 16 as AI-enhanced, and 11 as AI-critical. In percentage terms, this corresponds to 35.71%, 38.10%, and 26.19%, respectively (Table 2).

AI-dependent episodes were characterized by direct reliance on AI-generated outputs with minimal learner elaboration or verification. AI-enhanced episodes involved partial integration of AI support into student reasoning, with observable processes of selection, reformulation, or contextual adaptation. AI-critical episodes reflected the highest level of interpretive autonomy, as learners triangulated AI responses, questioned their adequacy, contextualized them, or used them dialogically rather than deferentially.

This distribution indicates that while AI-supported and hybrid forms of reasoning are already prevalent, fully critical engagement remains less frequent. The observed pattern suggests a transitional learning ecology in which interpretive support is widespread, but the development of epistemic autonomy is still uneven. These trajectories were not predefined categories, but abductively derived from the observed episode patterns and subsequently aligned with the theoretical framework.

Table 2

Distribution of Interactional Episodes Across Developmental Trajectories

Developmental trajectory	Frequency (n)	Percentage (%)
AI-dependent interpretation	15	35.71%
AI-enhanced interpretation	16	38.10%
AI-critical interpretation	11	26.19%
Total	**42**	**100%**

Figure 3

Distribution of Developmental Trajectories in AI-Mediated Interpretation

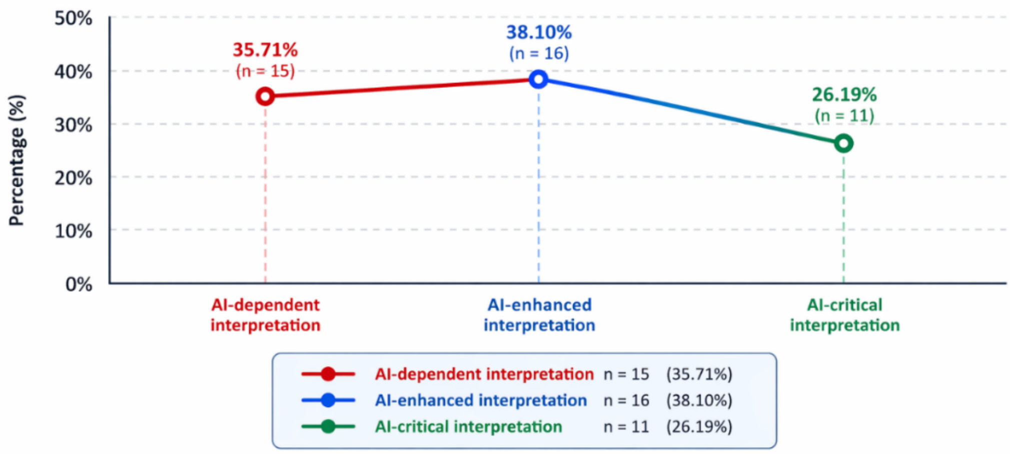


Figure 3 presents the distribution of the three developmental trajectories identified in the observational corpus: AI-dependent interpretation (n = 15), AI-enhanced interpretation (n = 16), and AI-critical interpretation (n = 11). These categories reflect increasing levels of interpretive autonomy and epistemic responsibility. The figure shows that hybrid and partially scaffolded AI use is the most frequent pattern, while fully critical engagement, although clearly present, remains less common. This empirical distribution supports the argument that pedagogical intervention is necessary to move learners from assisted use toward critically regulated interpretation. These trajectory distributions were abductively derived from the same observational corpus and reflect the pattern consolidation described in Sections 3.3 and 3.6.

5. Discussion

The findings of this study indicate that artificial intelligence functions not merely as a supportive tool but as a reconfigurator of interpretive cognition. By accelerating interpretive processes, AI compresses tasks that traditionally required sustained cognitive effort into immediate outputs, altering both the temporal structure and the depth of engagement in learning (Zhai et al., 2024). This shift aligns with broader cognitive concerns regarding the dominance of fast, intuitive processing over reflective reasoning (Kahneman, 2011).

From a hermeneutic perspective, this challenges the assumption that understanding emerges through dialogic engagement with texts and contexts (Gadamer, 2004). When learners rely on AI-generated outputs, interpretation risks becoming a process of epistemic consumption rather than production, echoing concerns in critical pedagogy (Freire, 1970) and broader debates on technology in education, where the tension between efficiency and depth remains central (Selwyn, 2023). At the same time, the findings show that AI can support higher-order thinking when used critically, and the developmental-critical model offers a structured framework for guiding learners toward interpretive autonomy through reflection, evaluation, and epistemic responsibility (Giannakos et al., 2025).

5.1. AI as a Reconfigurator of Interpretive Cognition

By enabling rapid access to structured interpretations, AI compresses complex reasoning into immediate outputs, privileging efficiency over depth (Zhai et al., 2024).

Although such acceleration enhances accessibility, it also introduces significant pedagogical risks. Interpretive reasoning in SSH disciplines depends on sustained engagement with ambiguity, uncertainty, and contextual complexity (Ricoeur, 1981), and when AI substitutes for these processes, learners may generate plausible interpretations without the cognitive restructuring necessary for deeper understanding (Karmiloff-Smith, 1992). Over time, this may lead to interpretive atrophy and reduced cognitive depth (Carr, 2010), making it essential to preserve interpretive effort in pedagogical design.

5.2. Hybrid Authorship and Argumentation

The hybridization of authorship identified in the results requires a reconsideration of writing within SSH education.

One consequence is the normalization of rhetorical patterns, as AI-generated texts often follow statistically dominant structures, producing coherent but frequently unoriginal arguments (Floridi, 2023). As students rely on such outputs, their writing may converge toward standardized forms, limiting diversity of expression and reinforcing concerns about homogenization in algorithmically mediated environments (Selwyn, 2023). Moreover, the uncritical integration of AI-generated evidence, including fabricated or unverifiable references (Zhai et al., 2024), contributes to the erosion of authorial voice and scholarly identity.

5.3. Distributed Agency and Epistemic Responsibility

The emergence of distributed epistemic agency marks a significant shift in knowledge production within AI-mediated learning environments. Drawing on distributed cognition theory, cognition extends across networks that include technological systems (Hutchins, 1995), with AI functioning as an epistemic participant that influences how knowledge is generated and validated.

Students increasingly attribute authority to AI-generated outputs because of their fluency and responsiveness, reflecting the emergence of machine authority in educational contexts (Selwyn, 2023; Giannakos et al., 2025). This contributes to a reconfiguration of epistemic trust but also creates the risk of epistemic outsourcing when learners relinquish responsibility for evaluation (Zhai et al., 2024). This highlights the importance of pedagogical strategies that cultivate epistemic responsibility.

5.4. Developmental-Critical Model: Six Checkpoints

The developmental-critical model proposed in this study provides a structured framework for guiding learners from AI-dependent to AI-critical interpretation. It is grounded in cognitive development, hermeneutics, and critical pedagogy (Piaget, 1972; Vygotsky, 1978; Gadamer, 2004; Freire, 1970), and is operationalized through six pedagogical checkpoints that scaffold interpretive autonomy.

The model begins with interpretive grounding and then moves through contextual anchoring of AI outputs, critical examination of algorithmic structures, hybrid authorship transparency, reflexive epistemic positioning, and dialogic engagement. Together, these checkpoints help prevent uncritical reliance on AI, maintain awareness of its epistemic limitations (Floridi, 2023), and balance the affordances of AI with the demands of interpretive learning (Zhai et al., 2024).

Figure 4

A Developmental-Critical Model for AI-Mediated SSH Pedagogy

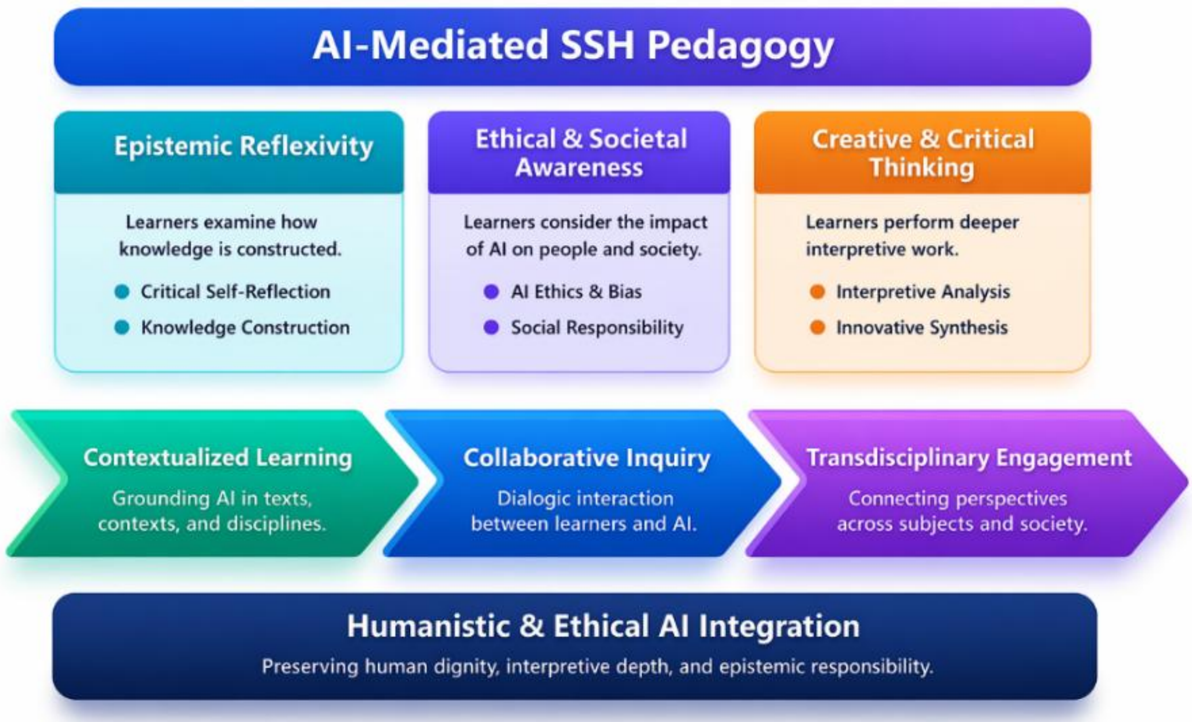


Figure 4 introduces a three-stage developmental-critical pedagogical model designed to scaffold interpretive autonomy in AI-rich SSH environments. Its stages include interpretive grounding prior to AI use, contextual anchoring of AI outputs, critical examination of algorithmic structures, transparency in hybrid authorship, reflexive epistemic positioning, and dialogic engagement under human control. The model is intended as a flexible framework whose application may vary across disciplines, learner profiles, and levels of AI integration. The six checkpoints of the developmental-critical model are derived directly from observed interactional patterns and correspond to transitions between the three developmental stages. Each checkpoint addresses a specific observed risk, including interpretive shortcutting, epistemic outsourcing, and loss of authorial voice, ensuring that pedagogical intervention is grounded in empirical classroom data rather than abstract theoretical assumptions.

5.5. Pedagogical Implications

The findings of this study carry significant implications for SSH pedagogy in AI-mediated contexts. First, interpretive grounding must be reinforced by ensuring that learners engage directly with primary materials before using AI tools, in line with hermeneutic principles emphasizing situated understanding and dialogic engagement (Gadamer, 2004).

Second, writing must be reconceptualized as a developmental rather than product-oriented activity, with tasks that prioritize reflection, iteration, and individual voice in accordance with constructivist and critical pedagogical traditions (Freire, 1970). Finally, critical AI literacy must be integrated across curricula, including technical, epistemic, and ethical competencies, while also supporting teacher preparedness for AI integration (Prilop et al., 2025; Alfarwan et al., 2025; Marzano, 2025; Dong, 2026).

6. Conclusion

The integration of generative artificial intelligence into Social Sciences and Humanities education represents a profound epistemic transformation extending beyond technological innovation. As AI systems increasingly participate in interpretive processes, they reshape how knowledge is produced, validated, and internalized by learners. This study has shown that AI functions as an epistemic technology (Floridi, 2023), fundamentally reconfiguring interpretive cognition through cognitive externalization, hybrid authorship, and distributed epistemic agency. These transformations challenge foundational assumptions of SSH education, especially the idea that understanding emerges through sustained dialogic engagement with texts and contexts (Gadamer, 2004; Ricoeur, 1981). The findings highlight risks such as interpretive shortcutting, epistemic outsourcing, and erosion of authorial voice, while also identifying opportunities for higher-order learning when AI is engaged critically (Zhai et al., 2024; Selwyn, 2023). The developmental-critical model offers a structured response grounded in cognitive-developmental theory (Piaget, 1972; Vygotsky, 1978), hermeneutics (Gadamer, 2004; Ricoeur, 1981), and critical pedagogy (Freire, 1970), providing a pathway for integrating AI into SSH education without sacrificing interpretive depth.

6.1. Limitations and Future Research

While the present study provides a theoretically grounded and empirically informed framework for understanding AI-mediated learning, it is based on qualitative, context-specific observations and does not aim for statistical generalization. Instead, it advances an analytically generalizable model grounded in established theoretical perspectives (Floridi, 2023; Selwyn, 2023). The findings are further limited by the use of a single classroom context and by the interpretive coding of naturally occurring interactions, which, while supporting analytical depth, do not permit claims of causal inference.

Accordingly, the applicability of the findings may vary across disciplinary contexts, learner characteristics, and levels of AI integration. Future research should empirically test the developmental-critical model across diverse educational settings and age groups. Quantitative and mixed-method approaches could offer additional insight into the impact of AI-mediated learning on cognitive development, interpretive depth, and epistemic agency, while longitudinal designs would be especially valuable for examining how sustained exposure to AI shapes critical thinking and interpretive autonomy over time (Zhai et al., 2024).

Further research is also needed on institutional and policy implications, including assessment practices, curriculum design, and teacher training, to ensure that AI integration supports rather than undermines the humanistic and interpretive foundations of SSH education (Gadamer, 2004; Freire, 1970).

References

Alfarwan, A. (2025). Generative AI use in K–12 education: A systematic review. *Frontiers in Education, 10*, 1647573. https://doi.org/10.3389/feduc.2025.1647573

Bobula, M. (2024). Generative artificial intelligence in higher education: Opportunities and challenges. *Journal of Learning Development in Higher Education, 29*, 1–15. https://doi.org/10.47408/jldhe.vi30.1137

Carr, N. (2010). *The shallows: What the Internet is doing to our brains*. W. W. Norton.

Dai, K., Liu, Y., & Zhang, X. (2026). Generative AI in higher education: A bibliometric review of emerging trends, power dynamics, and global research landscapes. *Computers & Education: Artificial Intelligence*, Article 100544. https://doi.org/10.1016/j.caeai.2026.100544

Daud, A., Aulia, A. F., Muryanti, M., Harfal, Z., Nabilla, O., & Ali, H. S. (2025). Integrating artificial intelligence into English language teaching: A systematic review. *European Journal of Educational Research, 14*(2), 677–691. https://doi.org/10.12973/eu-jer.14.2.677

Dong, Y. (2026). Generative AI technologies and educational outcomes: A comprehensive meta-analysis comparing traditional and AI-driven approaches. *Humanities and Social Sciences Communications*. Advance online publication. https://doi.org/10.1057/s41599-026-06903-y

Floridi, L. (2023). Artificial intelligence as a form of epistemic technology. *Philosophy & Technology, 36*(1), 1–18.

Freire, P. (1970). *Pedagogy of the oppressed*. Continuum.

Gadamer, H.-G. (2004). *Truth and method* (2nd rev. ed., J. Weinsheimer & D. G. Marshall, Trans.). Continuum. (Original work published 1975)

Garzón, J., Patiño, E., & Marulanda, C. (2025). Systematic review of artificial intelligence in education: Trends, benefits, and challenges. *Multimodal Technologies and Interaction, 9*(8), 84. https://doi.org/10.3390/mti9080084

Giannakos, M. N., Jaccheri, L., & Krogstie, J. (2025). The promise and challenges of generative AI in education. *Behaviour & Information Technology, 44*(11), 2518–2544. https://doi.org/10.1080/0144929X.2024.2394886

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Hutchins, E. (1995). *Cognition in the wild*. MIT Press.

Întorsureanu, I., Popescu, E., & Iordache, D. D. (2025). Generative AI in education: Perspectives through an academic lens. *Electronics, 14*(5), 1053. https://doi.org/10.3390/electronics14051053

Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers & Education: Artificial Intelligence, 8*, Article 100348. https://doi.org/10.1016/j.caeai.2024.100348

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. MIT Press.

Kuhn, D. (1991). *The skills of argument*. Cambridge University Press.

Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology, 50*(6), 2824–2838.

Marzano, D. (2025). Generative artificial intelligence (GAI) in teaching and learning processes at the K–12 level: A systematic review. *Technology, Knowledge and Learning*. Advance online publication. https://doi.org/10.1007/s10758-025-09853-7

Monib, W. K., Hasan, M. R., & Rahman, M. M. (2024). Generative AI and future education: A review, theoretical validation, and authors' perspective on challenges and solutions. *PeerJ Computer Science, 10*, e2105. https://doi.org/10.7717/peerj-cs.2105

Piaget, J. (1972). *The psychology of the child*. Basic Books.

Prilop, C. N., Mah, D.-K., Jacobsen, L. J., Hansen, R. R., Weber, K. E., & Hoya, F. (2025). Generative AI in teacher education: Educators' perceptions of transformative potentials and the triadic nature of AI literacy explored through AI-enhanced methods. *Computers & Education: Artificial Intelligence*, Article 100471. https://doi.org/10.1016/j.caeai.2025.100471

Ricoeur, P. (1981). *Hermeneutics and the human sciences: Essays on language, action and interpretation* (J. B. Thompson, Ed. & Trans.). Cambridge University Press.

Selwyn, N. (2023). *Education and technology: Key issues and debates* (3rd ed.). Bloomsbury.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications, 252*, 124167. https://doi.org/10.1016/j.eswa.2024.124167

Xiaoyu, W., Zhang, Y., Li, H., & Chen, X. (2025). Generative artificial intelligence in pedagogical practices: A systematic review of empirical studies (2022–2024). *Cogent Education, 12*(1). https://doi.org/10.1080/2331186X.2025.2485499

Zhai, X., Chu, X., Chai, C. S., Jong, M. S.-Y., Istenic, A., Spector, J. M., Liu, J.-B., Yuan, J., & Li, Y. (2024). Generative AI in education: Opportunities, challenges, and future research directions. *Computers & Education: Artificial Intelligence, 5*, 100118

AI Tools Declaration

AI-assisted tools were used only for minor linguistic editing (grammar, clarity, and style). Specifically, ChatGPT (OpenAI) was used solely for language refinement, without influencing the academic content, methodology, interpretation, or conclusions of the manuscript.

Author Biography

Cristina-Georgiana Voicu holds a PhD in Philology (British and American Cultural Studies) from Alexandru Ioan Cuza University of Iași. Her research focuses on the intersections of digital pedagogy, artificial intelligence in education, and cognitive and cultural studies. She has published in peer-reviewed national and international journals on topics such as AI-mediated learning, digital ethics, and inclusive pedagogical frameworks. She is actively involved in international educational and research networks, particularly within Erasmus+ initiatives, contributing to innovative approaches in AI-enhanced and sustainable education.