Int. j. adv. multidisc. res. stud. 2026; 6(2):479-488

Received: 26-01-2026
Accepted: 06-03-2026

International Journal of Advanced Multidisciplinary Research and Studies

ISSN: 2583-049X

From Dialogue to Explanation: A Framework for Analysing Engagement and Reasoning with Generative AI in Physics Learning

Konstantinos T Kotsis
Lab of Physics Education and Teaching, Department of Primary Education, University of Ioannina, Greece

DOI: https://doi.org/10.62225/2583049X.2026.6.2.5987
Corresponding Author: **Konstantinos T Kotsis**

Abstract
The rapid integration of generative artificial intelligence (AI) into science classrooms has created an urgent need for robust methodological tools to capture how learning unfolds during AI-mediated interactions. Existing re-search in AI-in-education has predominantly focused on learner perceptions or outcome-based measures, offering limited insight into the processes through which students engage with disciplinary ideas and construct scientific explanations. This article proposes a theoretically grounded methodological framework for analysing learning processes in AI-supported physics classrooms by integrating the ICAP model of cognitive engagement with the Claim–Evidence–Reasoning (CER) framework for scientific explanation. The framework operationalises ICAP for analysing student–AI dialogue and adapts CER to evaluate the epistemic structure and causal coherence of student-generated explanations produced in AI-mediated inquiry contexts. An illustrative application demonstrates how patterns of interactive, constructive, active, and passive engagement in dialogue can be systematically related to variations in the depth and integration of students' physics reasoning. By foregrounding interactional processes and epistemic quality rather than tool-specific effectiveness, the framework supports process-oriented, theory-informed research on generative AI in science education. The proposed ana-lytic architecture is intended as a portable toolkit that can facilitate comparability across studies, support methodological transparency, and contribute to the cumulative development of knowledge about learning processes in AI-mediated physics education.

Keywords: Generative Artificial Intelligence, Physics Education, Dialogic Learning, Conceptual Change, AI Literacy

Introduction

Generative artificial intelligence (GenAI) systems capable of sustained natural-language interaction are being adopted in educational settings at a pace that has outstripped the development of robust research instrumentation for analysing how learning unfolds when students work with these tools. Recent empirical work has documented that learners readily appropriate systems such as ChatGPT for help with disciplinary problem solving and explanation-building, including in physics contexts, and that students often value immediacy, conversational accessibility, and the perceived tutoring function of AI-generated feedback [1]. At the same time, scholarship has emphasised that meaningful educational use of GenAI depends not only on access to tools but also on learners' and teachers' capacity for AI literacy, including critical evaluation of AI outputs, prompt practices, and epistemic judgement about the status of machine-generated claims [2]. This emphasis on AI literacy connects with broader concerns about students' scientific literacy, which remains a critical foundation for effective engagement with scientific information and digital knowledge sources [3]. Parallel to these developments, research on pedagogical conversational agents indicates that dialogic design features---such as responsiveness, personalised feedback, and socio-affective cues---can shape how learners engage in explanation, reflection, and persistence, but also highlights wide variation in evaluation methods and the need for analytically precise frameworks that are portable across contexts [4]. Consequently, the field faces a methodological challenge: how to characterise and measure learning-relevant processes in AI-mediated dialogue without reducing such interactions to either surface indicators (e.g., time-on-task) or outcome-only measures (e.g., post-tests) that do not capture the dynamics of sense-making, revision, and reasoning that are central to physics learning.

This methodological gap is particularly salient in physics education, where conceptual progress depends on students' ability to coordinate intuitive ideas with formal principles and to construct explanations that connect evidence to theory [5]. GenAI systems introduce a new interactional layer into this process by producing fluent explanations, counterexamples, and prompts that may elicit student articulation, but they may also encourage premature closure when students accept authoritative-sounding responses without reconstructing the underlying causal structure. Learning analytics scholars have therefore argued that GenAI both generates new forms of learning data (notably, conversational traces) and requires new analytic approaches capable of distinguishing human contribution from tool contribution and of interpreting dialogue as evidence of cognitive engagement rather than as mere text production [6]. In parallel, recent work demonstrates that generative AI can be productively integrated into physics instruction by supporting the design of experiment worksheets and inquiry materials for teachers [7]. Related work in educational technology has similarly noted that educators perceive generative AI as reconfiguring assessment, feedback, and classroom roles, thereby increasing the importance of research designs that can analyse process data and interactional patterns in authentic learning settings [8]. Within this landscape, the absence of shared, theory-grounded coding and rubric systems for AI--student exchanges limits comparability across studies, complicates replication, and constrains the accumulation of evidence about when and why AI-supported inquiry leads to deeper reasoning rather than superficial performance.

The present article addresses this need by proposing and illustrating a methodological framework that integrates two complementary analytic lenses: the ICAP framework for cognitive engagement and the Claim--Evidence--Reasoning (CER) framework for scientific explanation. ICAP conceptualises engagement as a hierarchy of overt learning activities---Interactive, Constructive, Active, and Passive---and posits that the quality of cognitive processing increases when learners generate inferences and co-construct meaning through dialogue rather than merely receiving or manipulating information [9]. CER, in turn, offers a well-established structure for characterising the epistemic quality of students' explanations by examining whether claims are supported by appropriate evidence and linked through reasoning that invokes disciplinary principles [10]. Recent applied research continues to use and refine CER-oriented approaches to promote and assess scientific reasoning, underscoring its relevance as a practical rubric for evaluating explanation quality across grade levels [11]. However, neither ICAP nor CER was originally designed for contexts in which a generative system actively shapes the conversational environment by providing prompts, candidate explanations, and language models of scientific discourse. This paper therefore contributes a set of operational definitions and procedures that adapt ICAP to AI--student dialogue turns and adapt CER to student artifacts produced in AI-mediated inquiry, with attention to reliability and interpretive validity.

Accordingly, the purpose of this study is methodological rather than evaluative. Instead of asking which AI system is "better," the paper demonstrates how researchers can analyse AI-mediated learning processes in physics by (a) coding student--AI dialogue for engagement modes using an ICAP-informed scheme calibrated for human--AI turn-taking and (b) coding student explanations and related artifacts for reasoning structure using an adapted CER rubric sensitive to the provenance of evidence and the completeness of causal links. The study addresses the following methodological questions: how ICAP can be operationalised for AI--student discourse in physics inquiry, how CER can be adapted for AI-mediated student explanations, what analytic value is gained by combining these lenses, and what reliability and validity challenges arise when conversational traces and written artifacts are treated as complementary evidence of learning processes. By offering a replicable analytic workflow, explicit coding rules, and a rationale for triangulating engagement and reasoning measures, the paper aims to support cumulative research on GenAI in science education and to enable more precise claims about how dialogic interaction and explanation quality relate within AI-supported physics learning environments.

Conceptual and Analytical Foundations

The increasing integration of generative artificial intelligence into educational environments has intensified scholarly attention to how learning should be conceptualised and studied in technology-mediated contexts. Contemporary perspectives in learning sciences emphasise learning as a dynamic, process-oriented phenomenon that unfolds through interaction, sense-making, and iterative refinement of understanding rather than as a static outcome measured solely through pre- and post-test performance. This process-oriented view is particularly salient in physics education, where conceptual development involves the progressive coordination of intuitive reasoning with formal scientific representations and principles. In AI-supported learning environments, this coordination is mediated not only by human interlocutors but also by algorithmic systems that participate in discourse, provide explanations, and structure the interactional space. Recent work in educational technology has argued that generative AI tools fundamentally reshape the epistemic conditions of learning by introducing new forms of dialogic scaffolding, feedback, and representational support. For example, Vakarou *et al.* [12] document how AI can be operationalised through practical tools and lesson plans in physics classrooms, underscoring the need for analytic approaches that foreground interactional processes and epistemic practices rather than only learning products [13].

Within science education, dialogue has long been recognised as a central mechanism through which learners externalise ideas, negotiate meaning, and appropriate disciplinary ways of reasoning. The analysis of discourse provides access to the epistemic work of learners as they articulate claims, confront anomalies, and integrate evidence with theory. In AI-mediated contexts, dialogue acquires an additional layer of complexity, as students engage in sense-making not only with peers and teachers but also with artificial agents that generate scientifically framed responses and prompts. Research on pedagogical conversational agents suggests that dialogic interaction with artificial interlocutors can elicit explanation, reflection, and sustained engagement, but also highlights the risk that students may treat AI-generated responses as authoritative knowledge claims, thereby short-circuiting productive struggle and conceptual reconstruction [4]. From this perspective, AI functions as both a cognitive tool and an epistemic mediator, shaping the form and content of student reasoning. Analytical frameworks for studying learning in such environments must therefore be sensitive to how interactional patterns emerge in human--AI dialogue and how these patterns relate to deeper processes of conceptual understanding in physics.

The ICAP framework offers a theoretically grounded model for categorising observable forms of cognitive engagement based on the generativity and interactivity of learners' overt behaviours. By distinguishing Interactive, Constructive, Active, and Passive modes of engagement, ICAP provides a lens for linking observable interactional patterns to hypothesised differences in underlying cognitive processing, with interactive and constructive activities associated with deeper learning outcomes than merely active or passive engagement. Although ICAP was originally formulated for human--human and human--content interactions, recent studies have begun to explore its applicability to technology-mediated and dialogic learning environments, suggesting that it can serve as a useful heuristic for analysing engagement in AI-supported contexts when operational definitions are carefully adapted [9, 14]. However, the presence of generative AI introduces methodological challenges for ICAP-based analysis, as the co-construction of meaning occurs across human and artificial contributions, complicating the attribution of generativity and interaction to the learner alone. This necessitates a refined application of ICAP that explicitly accounts for the distinctive turn-taking structures and epistemic roles characteristic of AI-mediated dialogue.

Complementing engagement-focused analyses, the Claim--Evidence--Reasoning framework provides a widely used structure for examining the epistemic quality of scientific explanations produced by learners. CER foregrounds the coordination of claims with empirical or theoretical evidence and the articulation of reasoning that links evidence to claims through disciplinary principles. In physics education research, CER-based analyses have been employed to characterise the development of explanatory competence and to support instructional designs aimed at strengthening causal reasoning. Recent empirical studies continue to demonstrate the value of CER as an analytic rubric for capturing variations in explanation quality across educational levels and instructional contexts, particularly in inquiry-oriented science learning environments [11, 15]. In AI-supported settings, however, the provenance of evidence and the sources of reasoning become less transparent, as students may appropriate AI-generated information within their explanations. This raises methodological questions about how to code and interpret CER components when elements of evidence and reasoning are partially mediated by an artificial agent rather than constructed solely through students' own empirical activity.

The integration of ICAP and CER within a unified analytic framework is motivated by their complementary foci on engagement processes and epistemic structure, respectively, and aligns with broader efforts in physics education research to develop theory-informed design frameworks for technology-supported learning environments [16]. ICAP provides insight into how students participate in learning activities and interact with AI systems at the level of observable discourse, while CER offers a means of assessing the coherence and disciplinary alignment of the explanations that students ultimately produce. Together, these lenses enable a multi-dimensional analysis of AI-mediated learning in physics that captures both the dynamics of interaction and the quality of conceptual reasoning. Recent calls in the learning analytics and AI-in-education literature have underscored the need for such integrative frameworks that can connect interactional traces to epistemic outcomes, thereby supporting more nuanced interpretations of how generative AI shapes learning processes in complex classroom ecologies [6, 8]. By situating ICAP and CER within a coherent methodological architecture, the present study responds to these calls and provides a theoretically grounded basis for analysing how engagement patterns in AI--student dialogue relate to the structure and depth of students' physics explanations.

Research Context and Data Sources

The methodological framework advanced in this study was developed and illustrated within the context of AI-supported inquiry activities in secondary physics education. The research context reflects contemporary classroom conditions in which students increasingly encounter generative artificial intelligence as a readily available cognitive and epistemic resource. Recent studies indicate that secondary and post-secondary learners already appropriate generative AI tools for explanation-seeking, problem solving, and language-mediated sense-making in STEM domains, often in ways that are only partially aligned with instructional intentions [1, 6]. At the same time, research has shown that the level of scientific literacy among pre-service primary teachers plays a critical role in how effectively scientific concepts and technological tools are integrated into classroom practice [17]. Complementing learner-focused evidence, teacher perspectives highlight both the pedagogical opportunities and practical constraints of integrating ChatGPT into STEM classrooms, particularly for supporting inquiry-oriented activities and scaffolding explanation [18]. This growing presence of AI in everyday learning practices underscores the importance of situating methodological development within authentic classroom environments rather than controlled laboratory settings, as the interactional norms, task structures, and epistemic expectations of school physics classrooms shape how students engage with AI-mediated dialogue and how learning processes can be meaningfully analysed. The inquiry-based physics context adopted here foregrounds conceptual reasoning about phenomena such as motion and energy, domains in which students' intuitive ideas frequently diverge from formal scientific models and where dialogue, experimentation, and explanation are central to conceptual development [19]. Experimental activities that connect abstract physical constants with observable laboratory procedures can significantly support students' conceptual understanding. For example, the experimental determination of the Avogadro constant through classroom-based investigations provides students with opportunities to link microscopic concepts with measurable macroscopic quantities [20]. In addition experimental demonstrations and simple classroom apparatus can play an important role in supporting conceptual understanding by making abstract thermal and physical processes observable to students [21].

The data sources employed in this methodological study were selected to capture complementary dimensions of learning processes in AI-mediated environments. First, transcripts of student--AI dialogue provide access to the microgenetic unfolding of sense-making, including questioning, hypothesis testing, and revision of ideas. Conversational traces of this kind have been identified in recent learning analytics research as a rich but underutilised source of evidence for studying engagement and epistemic practices in technology-supported learning, particularly in contexts involving generative AI systems that participate actively in discourse [6]. Second, student-generated explanatory artifacts, including written explanations and representational products such as concept maps, offer insight into how interactional processes are recontextualised and stabilised in more durable forms of reasoning. Prior work in science education has demonstrated that written explanations provide a window into students' integration of claims, evidence, and reasoning, and that such artifacts are sensitive to instructional scaffolds and dialogic supports [15]. By analysing both dialogue and artifacts, the present methodological approach aligns with calls for multi-source data strategies that connect interactional processes with epistemic outcomes in AI-supported learning environments [8].

Ethical and methodological considerations are central to research that analyses fine-grained interactional data involving minors and AI systems. The collection and analysis of dialogue transcripts necessitate careful procedures for anonymisation, informed consent, and responsible data stewardship, particularly given the potential sensitivity of conversational content and the evolving regulatory landscape surrounding educational data and AI use. Recent policy-oriented scholarship has highlighted the importance of embedding ethical safeguards into AI-in-education research designs, including transparency about data use, protection of learner privacy, and critical reflection on the epistemic authority attributed to AI-generated content [22]. In methodological terms, the contextual specificity of classroom-based data also constrains claims of statistical generalisability, reinforcing the need to frame findings in terms of analytic generalisation and theoretical transferability rather than population-level inference. The present study therefore positions generative AI not as an instructional substitute but as a mediating artifact embedded within teacher-guided inquiry practices, including the planning and enactment of physics classroom experiments supported by AI-based tools [23], and treats the research context as an illustrative site for developing and refining analytic tools that can be adapted and validated across diverse educational settings and AI platforms.

Methodological Framework

The methodological framework proposed in this study integrates two complementary analytic lenses, the ICAP framework for cognitive engagement and the Claim--Evidence--Reasoning (CER) framework for scientific explanation, into a coherent procedure for analysing learning processes in AI-mediated physics classrooms. The framework is designed to be portable across classroom contexts and generative AI platforms, while remaining theoretically grounded in learning sciences and science education research. Its central aim is to provide researchers with a systematic way to move from raw interactional traces and student-produced artifacts to analytically meaningful characterisations of engagement and reasoning that can support cumulative knowledge building in the emerging field of AI-supported learning. Recent methodological discussions in learning analytics and educational technology have emphasised the need for transparent, theory-informed analytic pipelines that can handle fine-grained interactional data generated in digitally mediated learning environments, including conversational logs produced through interaction with generative AI systems [6, 24].

The first component of the framework involves the operationalisation of the ICAP model for AI--student dialogue. ICAP conceptualises learning-relevant activity in terms of four hierarchically ordered modes of engagement---Interactive, Constructive, Active, and Passive---each associated with qualitatively different forms of cognitive processing and learning potential. While ICAP has been widely applied in analyses of classroom interaction and technology-enhanced learning, the extension of this framework to human--AI dialogue requires careful adaptation because generative AI systems contribute substantively to the interactional sequence. In the present framework, the unit of analysis is defined as the student's conversational turn, with engagement categories assigned based on the epistemic work performed by the student in relation to the AI's preceding contribution. Interactive engagement is operationalised as instances in which students build on, challenge, or revise ideas in response to AI prompts or feedback, thereby participating in a co-construction of meaning. Constructive engagement is identified when students generate new inferences, explanations, or hypotheses without explicit uptake of the AI's prior utterance. Active engagement is coded when students manipulate information or apply procedures, such as substituting values into formulas or restating provided explanations, without generating new conceptual content. Passive engagement is characterised by minimal transformation of information, such as acknowledgements or verbatim acceptance of AI-generated statements. This operationalisation aligns with the theoretical commitments of ICAP while accommodating the distinctive turn-taking structures of AI-mediated dialogue, in which the interlocutor is a probabilistic language model rather than a human partner [9, 14]. The explicit focus on the student's epistemic contribution addresses recent methodological concerns about attributing cognitive engagement in settings where algorithmic systems actively shape the interactional environment [6].

The second component of the framework adapts the CER model to analyse the epistemic quality of student-produced artifacts generated in AI-supported inquiry contexts. CER has been extensively used to characterise scientific explanation in school science, with a substantial body of research demonstrating its utility for capturing the coherence and disciplinary alignment of students' reasoning. In AI-mediated settings, however, the provenance of claims, evidence, and reasoning becomes more complex, as students may appropriate AI-generated information within their explanations. The adapted CER procedure therefore treats claims as explicit explanatory assertions articulated by students, evidence as references to data, observations, or authoritative information sources invoked in support of claims, and reasoning as the explicit articulation of disciplinary principles that connect evidence to claims. To capture variation in explanatory depth, the framework employs an ordinal rubric distinguishing descriptive or procedural explanations, partial causal explanations, and integrated causal explanations that coherently link multiple variables or principles. This adaptation reflects recent refinements of CER-oriented assessment that emphasise the importance of causal integration and mechanistic reasoning in evaluating explanation quality in science learning [11, 15]. By foregrounding the structure of reasoning rather than the surface correctness of statements, the framework aims to distinguish between explanations that merely reproduce authoritative language and those that reflect deeper conceptual integration, a distinction that is particularly salient in the context of generative AI, where fluent but shallow reproductions of scientific discourse are readily available to learners.

The integration of ICAP and CER within a unified analytic procedure constitutes the third component of the framework. Analytically, dialogue transcripts are first segmented into student turns and coded for engagement mode using the adapted ICAP definitions, while student artifacts are independently coded for explanation quality using the CER-based rubric. The two analytic streams are then brought into relation through cross-mapping procedures that examine patterns between modes of engagement observed in dialogue and the epistemic structure of subsequent explanations. This triangulation logic is grounded in the assumption, supported by learning sciences research, that engagement processes and reasoning quality are related but not reducible to one another, and that meaningful inferences about learning require attention to both interactional dynamics and epistemic outcomes [9, 24]. Recent work in multimodal learning analytics similarly argues that combining process-oriented interaction data with artifact-based measures of reasoning provides a more comprehensive picture of learning than either data source alone, particularly in complex, AI-mediated learning ecologies [6].

Overall, the methodological framework articulated here responds to calls for theory-driven, replicable approaches to analysing learning in generative AI contexts. By specifying units of analysis, operational definitions, and an explicit workflow for integrating engagement and reasoning measures, the framework seeks to support methodological transparency and comparability across studies. At the same time, it remains sensitive to the distinctive epistemic conditions introduced by AI-mediated dialogue, in which the boundaries between learner-generated and tool-generated contributions are porous. In this respect, the framework is intended not as a fixed coding scheme but as a principled analytic architecture that can be iteratively refined as research on AI-supported learning matures and as generative AI systems continue to evolve.

Reliability, Validity, and Analytic Challenges

The credibility of analytic inferences drawn from AI-mediated learning data depends critically on the reliability and validity of the coding procedures used to characterise engagement and reasoning. In qualitative and mixed-methods research on learning processes, reliability is not merely a statistical property but a function of the transparency, stability, and interpretability of analytic categories across coders and contexts. Recent methodological work in learning analytics and educational research emphasises that the increasing availability of fine-grained digital trace data, including conversational logs generated through interaction with generative AI, necessitates explicit reliability protocols to ensure that analytic judgments about engagement and reasoning are not idiosyncratic or overly sensitive to contextual variation [24]. Within the present framework, reliability is addressed through the use of clearly specified operational definitions for ICAP engagement modes and CER components, coder training based on shared exemplars, and iterative refinement of coding rules through negotiated agreement. Such procedures align with best practices in qualitative content analysis and discourse-oriented learning analytics, which stress the importance of documenting analytic decisions and providing sufficient methodological detail to enable replication and secondary analysis [6, 25].

Validity considerations in the analysis of AI--student dialogue extend beyond conventional concerns about construct representation to encompass the epistemic status of AI-generated contributions and their influence on student discourse. Construct validity requires that ICAP codes meaningfully capture differences in cognitive engagement rather than superficial variations in linguistic form, and that CER-based ratings reflect the epistemic coherence of students' explanations rather than the fluency of language borrowed from AI outputs. Recent critiques of generative AI in education caution that students may appropriate authoritative-sounding explanations without fully internalising the underlying conceptual relations, thereby producing artifacts that appear epistemically sophisticated while masking shallow understanding [26]. This risk underscores the need for analytic criteria that prioritise causal integration, mechanistic coherence, and the explicit articulation of reasoning over mere terminological accuracy. Interpretive validity further requires sensitivity to the interactional context in which utterances are produced, as students' engagement modes may be shaped by the affordances and constraints of the AI interface, the framing of tasks by teachers, and the norms of classroom participation. Methodologically, this calls for analytic triangulation across dialogue, artifacts, and contextual data, in line with contemporary recommendations for enhancing the interpretive robustness of learning analytics in complex educational ecologies [24].

Beyond issues of reliability and validity, the analysis of AI-mediated dialogue introduces distinctive methodological challenges that complicate straightforward application of existing frameworks. One such challenge concerns the attribution of epistemic agency in human--AI interaction. In dialogic sequences, the generativity of student contributions is often co-constituted by AI prompts that scaffold particular forms of reasoning or invite specific linguistic structures. As a result, distinguishing between genuinely constructive engagement and AI-elicited performance becomes analytically delicate. Relatedly, the turn-taking structures of AI-mediated dialogue differ from those of human conversation, with AI systems capable of producing extended, highly structured responses that may reframe the problem space and constrain subsequent student moves. Recent work in human--computer interaction and AI-in-education highlights that such interactional asymmetries can shape learner behaviour in ways that are not readily captured by analytic models developed for human--human discourse [6, 22].

Another analytic challenge arises from the performative dimension of student discourse in AI-mediated environments. Students may adopt epistemic language, such as causal connectives or disciplinary terminology, in response to perceived expectations of the AI system, without necessarily engaging in the underlying conceptual work that such language conventionally signals. This phenomenon, which resonates with earlier critiques of "ritualised" scientific language use in classroom discourse, is potentially amplified by generative AI systems that model polished scientific explanations and thereby set implicit norms for acceptable responses. From a methodological standpoint, this reinforces the importance of combining engagement-based analyses with artifact-based reasoning assessments, as well as the need for caution in interpreting linguistic sophistication as a proxy for conceptual depth [15, 26]. Finally, the evolving nature of generative AI platforms introduces an additional layer of methodological instability, as changes in model behaviour, interface design, and default prompting practices can alter the interactional conditions under which data are generated. This temporal variability complicates longitudinal comparability and underscores the need for analytic frameworks that are sufficiently abstract to remain applicable across different AI systems while remaining sensitive to contextual specificities.

Taken together, these considerations indicate that methodological rigor in AI-mediated learning research cannot be achieved solely through the transplantation of established coding schemes into new technological contexts. Instead, it requires an explicit theorisation of the epistemic and interactional conditions introduced by generative AI and a corresponding adaptation of analytic practices. The framework advanced in this study addresses these challenges by foregrounding operational clarity, triangulation across data sources, and reflexive attention to the interpretive limits of engagement and reasoning codes in AI-supported inquiry. In doing so, it contributes to the development of a methodological repertoire capable of supporting cumulative, theory-informed research on learning processes in the rapidly evolving landscape of AI in physics education.

Illustrative Application of the Framework (Mini Case)

To demonstrate the analytic affordances of the proposed methodological framework, this section presents an illustrative application of the integrated ICAP--CER approach to a small subset of AI-mediated physics learning interactions. The purpose of this illustrative case is not to provide evaluative claims about instructional effectiveness, but to show concretely how engagement modes in AI--student dialogue can be systematically related to the epistemic structure of students' subsequent explanations. Recent methodological discussions in learning analytics stress the importance of such worked examples for making analytic frameworks transparent and usable by other researchers, particularly when dealing with complex, multimodal data generated in digitally mediated learning environments [6, 24]. By tracing the analytic steps from raw dialogue to coded engagement categories and from student artifacts to CER-based reasoning levels, the present mini case exemplifies how the framework can be operationalised in practice.

In the illustrative interaction, a student engages in a sequence of exchanges with a generative AI system while working on a conceptual physics task involving free-fall motion. Consistent with the view that alternative ideas in physics function as enduring cognitive resources rather than mere misconceptions [5], the student initially articulates an intuitive claim about the influence of mass on falling speed. Research in physics education has repeatedly shown that university students often hold persistent misconceptions about fundamental concepts of classical mechanics, particularly regarding force, motion, and acceleration [27]. The student then encounters an AI response that prompts reconsideration of this claim in light of Newtonian principles. Subsequent student turns include requests for clarification, reformulations of the original claim, and tentative integration of formal reasoning into the emerging explanation. When coded using the adapted ICAP scheme, these turns can be distinguished in terms of engagement mode, with instances of interactive engagement identified when the student explicitly builds on and revises ideas in response to AI prompts, and constructive engagement identified when the student generates new inferences without direct uptake of the preceding AI utterance. This differentiation illustrates how the framework captures variation in the epistemic work performed by the student across turns, rather than treating the dialogue as a homogeneous interactional stream. Such fine-grained differentiation is consistent with prior research demonstrating that shifts between engagement modes within a learning episode are associated with qualitatively different forms of cognitive processing and learning potential [9, 14].

The same learning episode can be examined through the lens of CER by analysing the student's written explanation produced after the dialogue. In the illustrative case, the explanation contains an explicit claim regarding the constancy of acceleration in free fall, references to gravitational force as evidence, and varying degrees of reasoning that link these elements through Newton's second law. Applying the adapted CER rubric enables the categorisation of the explanation according to its level of causal integration, distinguishing between descriptive or procedural formulations and more fully integrated causal accounts. This analytic move highlights how the epistemic quality of the student's explanation reflects, but does not mechanically mirror, the engagement patterns observed in the preceding dialogue. Recent work on explanation in science education underscores that the production of coherent, causally integrated explanations is sensitive to the quality of prior sense-making activities, but also shaped by representational demands and task framing, reinforcing the value of examining dialogue and artifacts in relation rather than in isolation [11, 15].

Bringing the ICAP-coded dialogue and CER-coded artifact into relation through cross-mapping allows for analytic inferences about how patterns of engagement may be associated with differences in explanation quality. In the illustrative case, turns characterised by interactive engagement, in which the student negotiates meaning with the AI and revises initial intuitions, are followed by a written explanation exhibiting a higher degree of causal integration than explanations produced after predominantly active or passive engagement sequences. While such patterns cannot be generalised from a single case, they exemplify the type of relational analysis enabled by the framework and illustrate how it can support theory-informed interpretations of AI-mediated learning processes. This integrative perspective aligns with recent calls in learning analytics and AI-in-education research to move beyond isolated metrics of engagement or performance and toward relational analyses that connect interactional processes with epistemic outcomes [6, 24].

The mini case also foregrounds interpretive cautions that accompany the analysis of AI-mediated dialogue. In particular, the student's uptake of AI-generated explanations raises questions about the extent to which observed improvements in explanation quality reflect internalised conceptual change as opposed to the appropriation of authoritative language patterns. This ambiguity underscores the importance of treating the mini case as illustrative rather than confirmatory and of situating analytic claims within a broader evidentiary base that may include longitudinal data, independent assessments, or triangulation with teacher observations. By making these interpretive limits explicit, the framework supports responsible methodological use and avoids overstating the inferential reach of fine-grained discourse and artifact analyses in AI-supported learning environments.

Implications for AI-in-Education Research

The methodological framework articulated in this study carries several implications for research on artificial intelligence in education, particularly for work that seeks to move beyond evaluative comparisons of tools toward a more fine-grained understanding of how learning unfolds in AI-mediated environments. One central implication concerns the status of interactional data as a legitimate and theoretically meaningful source of evidence about learning processes. The increasing availability of conversational logs generated through student interaction with generative AI systems creates new opportunities for process-oriented analysis, but it also demands analytic frameworks capable of interpreting such data in relation to established theories of engagement and reasoning. Recent scholarship in learning analytics has emphasised that the interpretive power of digital trace data depends on their integration with theoretically grounded constructs rather than on purely data-driven pattern detection [24]. By operationalising ICAP for AI--student dialogue and combining it with CER-based analysis of student explanations, the present framework provides a principled pathway for transforming conversational traces into theoretically interpretable indicators of engagement and epistemic quality, thereby contributing to the methodological maturation of AI-in-education research.

A second implication concerns the design of empirical studies in AI-supported learning contexts. Much of the early research on generative AI in education has focused on learner perceptions, usability, and short-term performance outcomes, reflecting both the novelty of these tools and the urgency of understanding their immediate impacts. However, recent reviews have cautioned that such outcome-centric approaches risk obscuring the mechanisms through which AI-mediated interaction shapes learning, motivation, and epistemic practices over time [8, 26]. The framework proposed here supports a shift toward process-sensitive research designs that foreground how students engage with AI prompts, how they negotiate meaning within human--AI dialogue, and how these engagement patterns relate to the structure of their scientific explanations. This orientation aligns with broader calls in the learning sciences for methodological approaches that can capture the temporal dynamics of learning in digitally mediated environments and that can inform the iterative design of AI-supported instructional interventions.

The framework also has implications for the emerging field of learning analytics, particularly in relation to the integration of qualitative and quantitative approaches. Learning analytics has traditionally privileged scalable, automated indicators derived from log data, clickstreams, and performance metrics, but recent work has highlighted the limitations of such indicators for capturing the epistemic and conceptual dimensions of learning, especially in open-ended inquiry tasks [24]. The ICAP--CER framework illustrates how fine-grained qualitative coding of dialogue and artifacts can complement computational analyses by providing theoretically interpretable categories that can, in principle, be operationalised for semi-automated or mixed-methods analytic pipelines. This opens avenues for future research that combines human-coded engagement and reasoning measures with machine learning approaches to discourse analysis, thereby advancing the methodological repertoire available for studying AI-mediated learning at scale while retaining sensitivity to epistemic nuance.

Finally, the methodological orientation advanced in this study has implications for comparative research across AI platforms and educational contexts. By focusing on engagement modes and reasoning structures rather than on the surface features of particular AI systems, the framework supports analytic generalisation across different generative models, interface designs, and instructional framings. This is particularly important given the rapid evolution of AI technologies and the corresponding instability of tool-specific affordances. Recent work in AI-in-education has underscored the need for research constructs that are robust to such technological change and that can support cumulative knowledge building despite shifting platforms and implementations [6]. In this respect, the ICAP--CER framework contributes to the development of a portable analytic vocabulary for studying AI-mediated learning processes, enabling researchers to compare findings across studies and contexts in terms of theoretically meaningful dimensions of engagement and epistemic quality rather than tool-specific performance metrics.

Limitations and Future Methodological Development

Despite the conceptual coherence and analytic promise of the ICAP--CER framework proposed in this study, several limitations must be acknowledged that delimit the scope of its current applicability and point toward directions for future methodological development. One important limitation concerns the reliance on textual interactional data and written student artifacts as primary sources of evidence for learning processes. While dialogue transcripts and explanatory texts provide rich access to aspects of engagement and reasoning, they capture only a subset of the multimodal resources through which students make sense of physics phenomena, including gestures, gaze, inscriptions, and interactions with physical or virtual representations. Recent research in multimodal learning analytics has demonstrated that these embodied and representational dimensions play a significant role in sense-making and conceptual coordination, particularly in science learning environments that involve diagrams, simulations, and hands-on activities [28]. The current framework does not explicitly incorporate such multimodal data streams, which constrains its capacity to account for non-verbal forms of engagement and may lead to an underestimation of learners' epistemic work when key aspects of reasoning are enacted outside of language. Future methodological extensions could integrate multimodal analytic techniques, including video-based interaction analysis and sensor-derived data, to provide a more comprehensive account of engagement and reasoning in AI-supported inquiry contexts.

A second limitation concerns the temporal scope of the analytic lens. The ICAP--CER framework, as articulated here, is primarily designed for episode-level analysis of engagement and reasoning within bounded learning activities. While this granularity is well suited to capturing microgenetic processes of sense-making in AI-mediated dialogue, it offers limited insight into the longitudinal development of conceptual understanding, epistemic beliefs, and learner identity over extended periods of AI use. Longitudinal perspectives are increasingly recognised as essential for evaluating the educational impact of generative AI, given concerns about novelty effects, shifting patterns of reliance on AI tools, and the potential reconfiguration of learners' epistemic practices over time [26, 29]. Future research could therefore extend the framework to incorporate longitudinal analytic designs, enabling researchers to trace trajectories of engagement modes and reasoning structures across sequences of AI-supported learning episodes and to examine how these trajectories relate to durable changes in conceptual understanding and scientific practices.

A further limitation pertains to the interpretive challenges introduced by the evolving nature of generative AI technologies. The behaviour of AI systems is shaped by ongoing model updates, interface redesigns, and changes in default prompting strategies, all of which can alter the interactional conditions under which learning data are generated. This technological volatility complicates methodological stability and raises questions about the comparability of findings across studies conducted with different versions of AI systems. Recent policy and research-oriented analyses of AI in education have underscored the need for methodological frameworks that are robust to such technological change and that foreground theoretically grounded constructs over tool-specific affordances [22]. While the ICAP--CER framework aims to provide such abstraction by focusing on engagement modes and reasoning structures, further validation across diverse AI platforms and instructional designs is required to establish its generalisability and to refine its operational definitions in light of emerging interactional patterns.

Finally, the current framework relies on human coding of dialogue and artifacts, which, while necessary for ensuring interpretive depth, limits scalability and raises questions about feasibility in large-scale studies. As AI-mediated learning environments generate increasingly large volumes of interactional data, there is a growing need for methodological approaches that can combine the interpretive richness of qualitative coding with the scalability of computational analysis. Emerging work in natural language processing and learning analytics suggests the possibility of semi-automated coding of discourse features related to engagement and reasoning, but such approaches remain methodologically immature and risk reifying surface linguistic patterns as proxies for cognitive processes [24]. Future methodological development could explore hybrid analytic pipelines in which machine learning techniques are trained on human-coded ICAP and CER data to support large-scale pattern detection while preserving theoretical alignment and interpretive oversight.

Taken together, these limitations underscore that the ICAP--CER framework should be understood as an initial methodological architecture rather than a finished analytic solution. Its further development will depend on iterative refinement through empirical application across diverse contexts, the incorporation of multimodal and longitudinal data, and the exploration of computational supports for scaling qualitative analysis. By situating these future directions within ongoing debates about the epistemic, ethical, and methodological challenges of AI in education, the framework invites continued methodological innovation aimed at capturing the complexity of learning processes in AI-mediated physics classrooms and beyond.

Conclusion

This article has advanced a methodological framework for analysing learning processes in AI-mediated physics education by integrating the ICAP model of cognitive engagement with the Claim--Evidence--Reasoning framework for scientific explanation. The central contribution of this work lies not in evaluating the effectiveness of particular generative AI systems, but in providing a theoretically grounded and replicable analytic architecture through which researchers can examine how students engage in dialogue with AI and how such engagement is reflected in the epistemic structure of their explanations. In the context of the rapid diffusion of generative AI in educational settings, the availability of robust methodological tools for interpreting interactional traces and student-produced artifacts is essential for moving the field beyond descriptive accounts of tool use toward cumulative, theory-informed research on learning processes.

By operationalising ICAP for AI--student dialogue, the framework foregrounds differences in the epistemic work performed by learners across interactional turns, distinguishing between interactive, constructive, active, and passive modes of engagement in a manner sensitive to the distinctive turn-taking structures of human--AI interaction. The adaptation of CER for AI-mediated student artifacts, in turn, provides a principled basis for examining the coherence and causal integration of students' physics explanations, addressing concerns that generative AI may facilitate the production of fluent but epistemically shallow responses. The integration of these lenses through cross-mapping procedures enables relational analyses that connect patterns of engagement in dialogue to the quality of subsequent reasoning, thereby supporting more nuanced interpretations of how AI-supported inquiry shapes conceptual sense-making in physics classrooms.

Beyond its immediate analytic utility, the framework contributes to ongoing debates about the epistemic and methodological challenges posed by generative AI in education. It responds to calls for process-oriented research designs that treat dialogue and explanation as primary data for understanding learning, and it aligns with broader efforts in learning analytics and educational technology to develop theory-informed approaches for interpreting complex digital trace data. At the same time, the framework remains open to refinement as AI technologies evolve and as empirical research accumulates across diverse instructional contexts. Its value lies in offering a shared analytic vocabulary and workflow that can support comparability across studies and facilitate the gradual construction of a cumulative evidence base on AI-mediated learning processes in science education.

In conclusion, the ICAP--CER framework is proposed as a portable methodological toolkit for researchers investigating generative AI in physics classrooms and related STEM domains. By enabling systematic analysis of engagement and reasoning in AI--student interaction, the framework contributes to the development of a more rigorous and process-sensitive research agenda in AI-in-education. Such an agenda is necessary if the field is to move beyond questions of whether generative AI "works" and toward a deeper understanding of how, when, and under what conditions AI-mediated dialogue can support meaningful conceptual learning in physics.

References

1. Ding L, Li T, Jiang S, Gapud A. Students' perceptions of using ChatGPT in a physics class as a virtual tutor. Int J Educ Technol High Educ. 2023; 20:63. Doi: 10.1186/s41239-023-00434-1
2. Walter Y. Embracing the future of artificial intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. Int J Educ Technol High Educ. 2024; 21:15. Doi: 10.1186/s41239-024-00448-3
3. Tsoumanis K, Stylos G, Kotsis KT. A comparative study between Greek pre-service teachers and primary school students' scientific literacy levels. Sci Educ Int. 2023; 34(2):121-131. Doi: 10.33828/sei.v34.i2.6
4. Ortega-Ochoa E, Arguedas M, Daradoumis T. Empathic pedagogical conversational agents: A systematic literature review. Br J Educ Technol, 2024. Doi: 10.1111/bjet.13413
5. Kotsis KT. Alternative ideas about concepts of physics are a timelessly valuable tool for physics education. Eurasia J Sci Environ Educ. 2023; 3(2):83-97. Doi: 10.30935/ejsee/13776
6. Yan L, Martinez-Maldonado R, Gašević D. Generative artificial intelligence in learning analytics: Contextualising opportunities and challenges through the learning analytics cycle. Proc ACM Hum-Comput Interact, 2024. Doi: 10.1145/3636555.3636856
7. Kotsis KT. ChatGPT develops physics experiment worksheets for primary education teachers. Eur J Educ Stud. 2024a; 11(5):1-20. Doi: 10.46827/ejes.v11i5.5274
8. Lee D, Arnold M, Srivastava A, Plastow K, Strelan P, Ploeckl F, *et al*. The impact of generative AI on higher education learning and teaching: A study of educators' perspectives. Comput Educ Artif Intell. 2024; 6:100221. Doi: 10.1016/j.caeai.2024.100221
9. Chi MTH, Wylie R. The ICAP framework: Linking cognitive engagement to active learning outcomes. Educ Psychol. 2014; 49(4):219-243. Doi: 10.1080/00461520.2014.965823
10. Krajcik J, McNeill KL. Designing and assessing scientific explanation tasks. In: Gunstone R, editor. Encyclopedia of science education. Springer, 2015. Doi: 10.1007/978-94-007-2150-0_48
11. Diola WY, Jalon JB Jr, Prudente MS. Exploring the use of claim-evidence-reasoning in promoting scientific reasoning skills of elementary school students. Anatol J Educ. 2025; 10(1):203-214. Doi: 10.29333/aje.2025.10115a
12. Vakarou G, Stylos G, Kotsis KT. AI for enhancing physics education: Practical tools and lesson plans. Int J Sci Math Technol Learn. 2024; 31(2):159-176. Doi: 10.18848/2327-7971/CGP/v31i02/159-176
13. Lampropoulos G, Papadakis S. The educational value of artificial intelligence and social robots. In: Lampropoulos G, Papadakis S, editors. Social robots in education. Studies in Computational Intelligence. Vol. 1194. Cham: Springer, 2025, 3-15. Doi: 10.1007/978-3-031-82915-4_1
14. Wekerle C, Daumiller M, Janke S, Dickhäuser O, Dresel M, Kollar I. Putting ICAP to the test: How technology-enhanced learning activities are related to cognitive and affective-motivational learning outcomes in higher education. Sci Rep. 2024; 14:16295. Doi: 10.1038/s41598-024-66069-y
15. McNeill KL, Berland LK. What is (or should be) scientific evidence use in K--12 classrooms? J Res Sci Teach. 2017; 54(5):672-689. Doi: 10.1002/tea.21381
16. Zourmpakis A-I, Kalogiannakis M, Papadakis S. A review of the literature for designing and developing a framework for adaptive gamification in physics education. In: Taşar MF, Heron PRL, editors. The International Handbook of Physics Education Research: Teaching Physics. AIP Publishing, 2023. Chapter 5. Doi: 10.1063/9780735425712_005
17. Stylos G, Siarka O, Kotsis KT. Assessing Greek pre-service primary teachers' scientific literacy. Eur J Sci Math Educ. 2023; 11(2):271-282. Doi: 10.30935/scimath/12637
18. Uğraş H, Uğraş M, Papadakis S, Kalogiannakis M. Innovative early childhood STEM education with ChatGPT: Teacher perspectives. Technol Knowl Learn. 2025; 30:809-831. Doi: 10.1007/s10758-024-09804-8
19. Kotsis KT. The significance of experiments in inquiry-based science teaching. Eur J Educ Pedagogy. 2024b; 5(2):86-92. Doi: 10.24018/ejedu.2024.5.2.815
20. Siafarikas M, Stylos G, Chatzimitakos T, Georgopoulos K, Kosmidis C, Kotsis KT. Experimental teaching of the Avogadro constant. Phys Educ. 2023; 58:065026. Doi: 10.1088/1361-6552/acfb6b
21. Stylos G, Kotsis KT. Use of a simple homemade apparatus to teach basic thermal concepts -- Six qualitative demonstrations/experiments. Phys Teach. 2021; 59:477-479. Doi: 10.1119/10.0006134
22. OECD. OECD Digital Education Outlook 2026: Exploring effective uses of generative AI in education. OECD Publishing, 2026. Doi: 10.1787/062a7394-en
23. Kotsis KT. Artificial intelligence helps primary school teachers to plan and execute physics classroom experiments. EIKI J Eff Teach Methods. 2024c; 2(2):1-9. Doi: 10.59652/jetm.v2i2.158
24. Raković M, Gašević D, Hassan SU, Ruipérez Valiente JA, Aljohani N, Milligan S. Learning analytics and assessment: Emerging research trends, promises and future opportunities. Br J Educ Technol. 2023; 54(1):10-18. Doi: 10.1111/bjet.13301
25. Hsieh H-F, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res. 2005; 15(9):1277-1288. Doi: 10.1177/1049732305276687
26. Kasneci E, Sessler K, Küchemann S, Bannert M, Dementieva D, Fischer F, *et al*. ChatGPT for good? On opportunities and challenges of large language models for education. Learn Individ Differ. 2023; 103:102274. Doi: 10.1016/j.lindif.2023.102274
27. Stylos G, Evangelakis GA, Kotsis KT. Misconceptions on classical mechanics by freshman university students: A case study in a physics department in Greece. Themes Sci Technol Educ. 2008; 1(2):157-177.
28. Ochoa X, Worsley M. Augmenting learning analytics with multimodal sensory data. J Learn Anal. 2016;

3(2):213-219. Doi: 10.18608/jla.2016.32.10
29. Zawacki-Richter O, Marín VI, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education. Int J Educ Technol High Educ. 2019; 16:39. Doi: 10.1186/s41239-019-0171-0