JOURNAL OF LEARNING ANALYTICS

SOLAR SOCIETY for LEARNING ANALYTICS RESEARCH

Volume (), . http://dx.doi.org/

EduHyperAgent: From Handwritten Learning Evidence to Multi-Agent Pedagogical Decision-Making

Anonymous author(s) – blinded manuscript

Abstract

The expansion of generative AI in education has increased the availability of specialised tools, yet coordinating them remains challenging. Handwritten STEM solutions capture intermediate steps, revisions, and explanations. However, recognising written content does not necessarily reveal the reasoning behind an error or determine an appropriate instructional response. Generative AI could connect this evidence with diagnostic dialogue and personalised activities under teacher oversight. Based on retrospective evidence from AI-supported educational feedback and the learning analytics cycle, this paper presents a methodology and the theoretical basis for creating EduHyperAgent: a multi-agent system governed by teachers for making adaptive pedagogical decisions. This methodology incorporates mathematical verification, optional oral diagnosis, and spline-based modelling of learning trajectories. A proposed decision algorithm uses spline-derived indicators to select between problem solving followed by instruction (PS-I) and instruction followed by problem solving (I-PS), subject to pedagogical constraints. We outline principles for coordinating specialised agents and identify Weighted Group Relative Policy Optimisation (GRPO) as a potential approach to future offline policy optimisation. This work establishes a conceptual and mathematical basis for system development. Implementation and prospective evaluation are required to determine its effects on independent learning, knowledge transfer, and teacher workload.

Notes for Practice

- The growing availability of generative AI tools does not necessarily improve learning outcomes. Effective integration requires coherent pedagogical strategies, reliable learning evidence, and teacher oversight of instructional decisions.
- This research is based on a retrospective five-year educational study in which approximately 1,000 students in grades 1–11 and 50 teachers took part of their own accord. The study's AI-supported feedback system informs ScriptaGrad, a specialised component for analysing handwritten STEM solutions.
- EduHyperAgent proposes integrating written assessment, oral diagnostic questioning, and personalised follow-up tasks within a teacher-governed multi-agent framework. Adaptive instructional pathways aim to support independent mastery; their effectiveness requires prospective validation.

Keywords

learning analytics; pedagogical agents; GRPO; reward weighting; multi-agent orchestration; handwritten STEM; ScriptaGrad; creative inquiry; human-AI collaboration

Submitted: — **Accepted:** — **Published:**

Affiliation withheld for double-blind review

1. Introduction

1.1 A longitudinal transition from testing to evidence-centred pedagogy

This study builds on a programme of pedagogical experiments initiated in Ukraine in 2019, when experimental classes began integrating artificial intelligence and digital tools into everyday teaching. Early applications included automated assessment, interactive visualisations, educational simulations, and rule-based learning systems. Their main purpose was to reduce routine

teacher workload and create more opportunities for creative, interactive, and personalised instruction.

These experiments also highlighted a key limitation of conventional teaching: although teachers use various educational resources and methods, classroom instruction and assessment often follow common procedures that do not fully capture individual learning processes. In particular, traditional testing provides limited information about how students develop ideas, justify solutions, recognise mistakes, and revise their reasoning. This challenge motivated the gradual transition from automated testing towards more detailed analysis of students' learning evidence and adaptive feedback. Our early experiments made a related assessment limitation apparent. Algorithmically scored multiple-choice items were efficient for administering and comparing large volumes of responses, but answer selection alone rarely disclosed the intermediate reasoning behind a decision. The same correct option could result from sound understanding, elimination, guessing, or external assistance; the same incorrect option could arise from very different cognitive and procedural difficulties. These limitations do not mean that multiple-choice assessment is intrinsically ineffective. They mean that it should not be the sole evidential basis for diagnosing open-ended mathematical reasoning.

An important methodological step towards richer assessment of mathematical reasoning was the integration of written problem-solving with oral explanation and diagnostic questioning. An early implementation of this approach was described in the context of the MRIIA+ educational platform, where learners were encouraged not only to provide written solutions but also to justify their reasoning and respond to follow-up questions (Zalizko, 2026a).

The public release of ChatGPT in November 2022 accelerated the adoption of generative AI in education (OpenAI, 2022). Our research increasingly focused on open-ended and creative tasks, allowing students to demonstrate their reasoning through written explanations, mathematical expressions, diagrams, and handwritten solutions. Unlike multiple-choice assessment, such tasks provide richer evidence of students' understanding, problem-solving strategies, and individual errors.

Due to the war in Ukraine, the research activities were relocated to Switzerland in February 2022, creating new opportunities to investigate multilingual learning and assessment in Ukrainian, German, and English. Meanwhile, the growing diversity of educational AI applications created new possibilities for supporting specific instructional tasks. Figure 1 presents a functional classification of AI-enabled educational tools used by teachers at the International Innovative School (IIS).

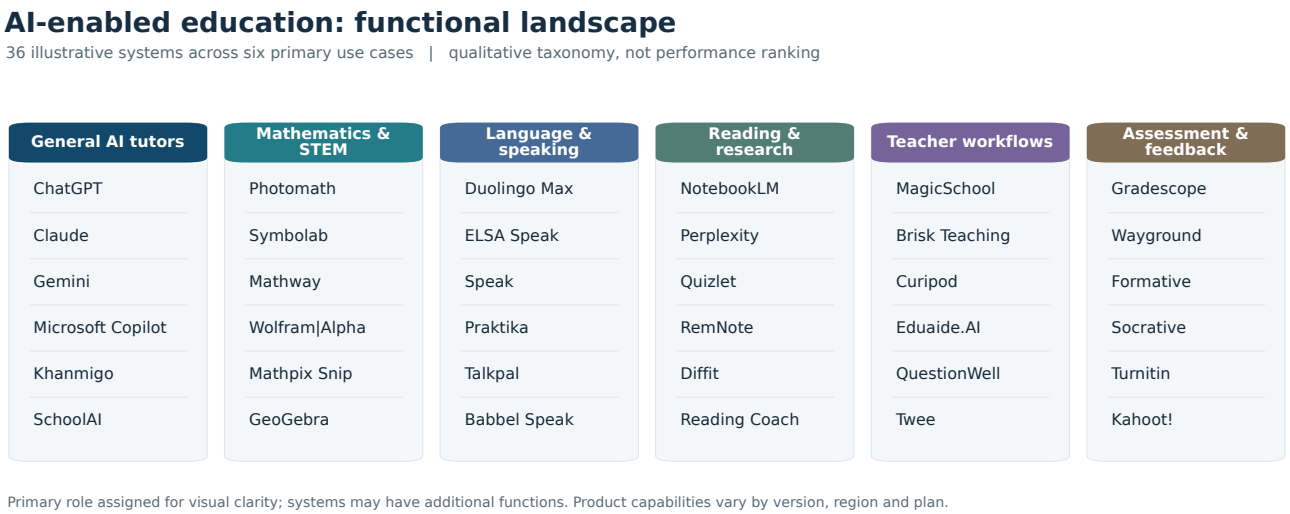


Figure 1. Systematization of AI-Enabled Educational Tools Used by Teachers at the International Innovative School.

Despite rapid improvements in multimodal recognition, important assessment gaps remain. Handwritten Cyrillic-language explanations, mixed scripts, corrections and unconventional mathematical symbols present challenges whose scale must be measured by language, script, task family and capture quality rather than assumed. More fundamentally, extracting text is not equivalent to interpreting a mathematical or physical argument. A handwritten thermodynamic solution, for example, may combine a p–V diagram with equations for heat, work and internal energy. In a Carnot cycle, the evaluator must distinguish isothermal from adiabatic transformations, interpret the direction of the cycle, verify sign conventions for $W = \oint p\,dV$, and assess whether the claimed efficiency $\eta = 1 - T_c/T_h$ is justified by an ideal reversible-engine model. A transcription that correctly recovers every formula can still miss a physically inconsistent process or an unjustified transition between steps. Similar difficulties arise in geometric proofs, chemical reaction pathways, and free-form explanations of experimental evidence.

Domain-aware verification and interpretation must therefore remain separate from transcription.

1.2 Why Error Detection Is Insufficient for Pedagogical Diagnosis

A central limitation of conventional assessment is that the correctness of a student's answer does not fully reveal the reasoning behind it. An incorrect solution may reflect a conceptual misunderstanding, a procedural slip, or an incomplete attempt, while a correct answer does not necessarily demonstrate deep understanding. Research on formative feedback and intelligent tutoring highlights the need to examine students' reasoning and provide support that responds to their learning needs (Hattie and Timperley, 2007; VanLehn, 2011).

This distinction motivates the ScriptaGrad approach to analysing intermediate steps in handwritten mathematical solutions and the proposed EduHyperAgent extension, which combines written evidence with targeted oral diagnostic questions.

This distinction may help explain a familiar educational paradox: students who consistently achieve high grades do not necessarily demonstrate superior creativity, independent reasoning, or success when facing unfamiliar problems. Conversely, students with lower grades may possess valuable problem-solving abilities that are not adequately captured by standardised assessment.

Consider a simple example from mathematics Figure 2. A student writes

$$3(x+2) = 3x+2, \tag{1}$$

where the correct expansion is $3x+6$. Although the error is mathematically clear, its underlying cause remains uncertain. The student may misunderstand the distributive property, have made a temporary calculation mistake, or have omitted a step despite otherwise understanding the concept. A single incorrect expression cannot reliably distinguish between these possibilities. Figure 2 shows two different levels of fraction structure in handwritten mathematical expressions, along with cancellation marks. While most contemporary AI models can interpret cancellation marks relatively easily, correctly identifying the main fraction bar and distinguishing it from sub-ordinate fraction bars remains challenging. This structural ambiguity is important because identifying the principal fraction determines the hierarchical relationships among mathematical components, and consequently the correctness of the reconstructed expression and its subsequent mathematical evaluation.

This creates an important distinction between error detection and pedagogical diagnosis. ScriptaGrad can support the identification and verification of errors in handwritten solutions, but understanding why an error occurred may require additional learning evidence.

EduHyperAgent is proposed to address this limitation through adaptive diagnostic interaction. Instead of immediately assigning a final judgement, the system may ask a short oral question, request an explanation, or generate a new written task targeting the possible misconception. Subsequent responses can help distinguish conceptual misunderstanding from procedural mistakes and guide the selection of appropriate learning support.

The central objective is therefore not simply to identify what a student answered incorrectly, but to determine what additional evidence is needed to support understanding and future progress.

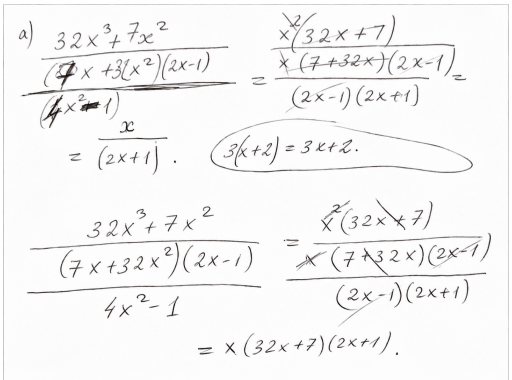


Figure 2. Excerpt from an answer to a maths test question.

Identifying a mathematical error is only the first step towards understanding a student's learning needs. A written solution may reveal what went wrong, but it rarely provides sufficient evidence to explain why the error occurred. We therefore propose that written STEM assessment should be complemented by short, targeted oral diagnostic questions. These questions can help

distinguish a conceptual misunderstanding from a temporary procedural mistake and provide additional evidence for selecting the next instructional action.

An important theoretical foundation for this approach is the distinction between two instructional sequences investigated by Kapur and colleagues: Instruction followed by Problem Solving (I-PS) and Problem Solving followed by Instruction (PS-I) (Kapur, 2014; Sinha and Kapur, 2021).

In the I-PS sequence, students first receive an explanation, demonstration, or worked example and then apply the acquired knowledge to a problem. This approach may be appropriate when students lack essential prerequisite knowledge or require explicit initial guidance.

In the PS-I sequence, students first attempt to solve a problem using their existing knowledge, explore possible strategies, and subsequently receive instruction that helps them compare, revise, and consolidate their solutions. A particular form of PS-I, known as Productive Failure, deliberately uses unsuccessful initial attempts as opportunities for deeper conceptual learning (Kapur et al., 2014; Sinha et al., 2021; Puech et al., 2025).

These instructional sequences provide two possible pathways for EduHyperAgent. Rather than applying either strategy uniformly, the proposed system uses written evidence and subsequent oral responses to inform the selection of an appropriate instructional sequence.

Consider the handwritten mathematical error illustrated in Figure 2. Rather than immediately providing the correct solution, EduHyperAgent initiates a short oral diagnostic interaction to explore the student's reasoning, clarify the possible causes of the error, and inform the selection of an appropriate instructional strategy.

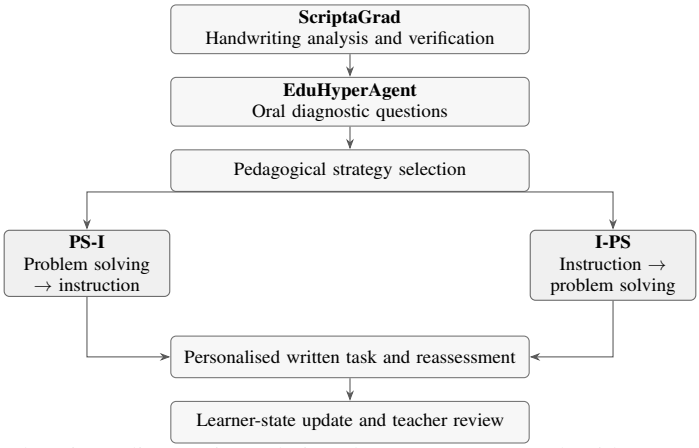


Figure 3. Proposed written–oral–written diagnostic cycle in EduHyperAgent. Oral evidence guides the selection of PS-I or I-PS instructional strategies and personalised reassessment.

The framework combines diagnostic questioning with adaptive instructional sequencing (Sinha and Kapur, 2021; Puech et al., 2025). Students may require different numbers of instructional cycles before demonstrating independent mastery. The effectiveness of this proposed approach remains to be evaluated experimentally.

1.3 Three Pedagogical Models: From Conventional Teaching to EduHyperAgent

Research on intelligent tutoring systems and formative feedback has demonstrated the importance of individualised instructional support and timely pedagogical interventions (VanLehn, 2011; Hattie and Timperley, 2007). More recent studies have highlighted the role of instructional sequencing, productive failure, and pedagogically guided AI tutoring (Sinha and Kapur, 2021; Puech et al., 2025).

Building on these foundations, EduHyperAgent proposes a transition from conventional instruction and general AI-assisted learning towards adaptive pedagogical orchestration. The framework coordinates specialised AI agents, integrates written and oral diagnostic evidence, and incorporates spline-based modelling of individual learning trajectories. These three conceptual approaches are compared in Figure 4.

1.4 From ScriptaGrad to the proposed EduHyperAgent

Between 2022 and 2025, our research increasingly focused on a fundamental challenge in AI-assisted STEM assessment: how to move beyond recognising handwritten answers towards interpreting and evaluating students' mathematical reasoning.

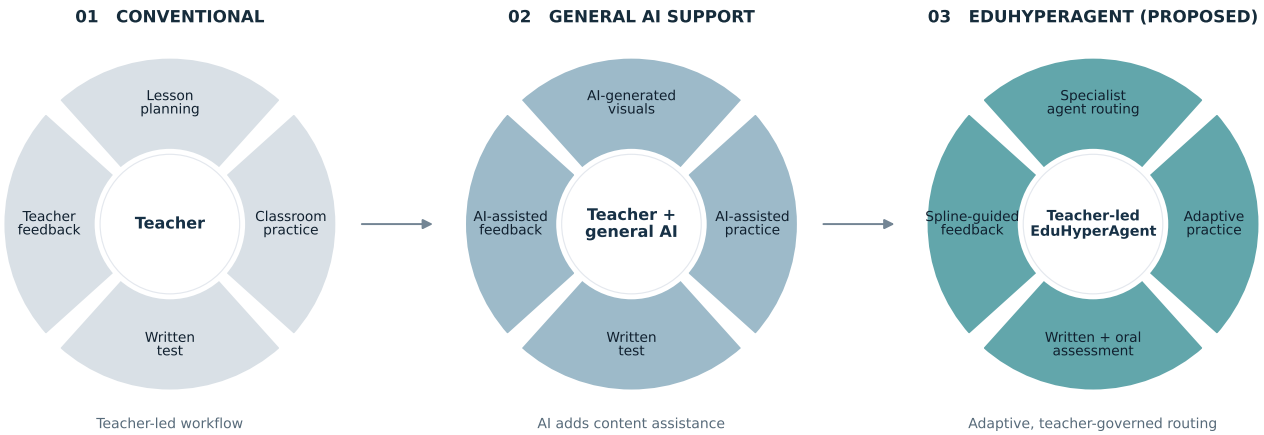


Figure 4. Evolution of pedagogical models: conventional instruction, general AI-assisted teaching, and the proposed EduHyperAgent framework.

The research addressed the detection, segmentation, transcription, and assessment of handwritten mathematics and physics solutions, including intermediate calculations, formulas, and diagrams. These investigations contributed to the development of *ScriptaGrad*, a specialised system for processing and assessing handwritten STEM work (Kortemeyer et al., 2025; ScriptaGrad, n.d.).

Building on research in automated mathematical assessment and multimodal learning analytics (Baral et al., 2023; Blikstein and Worsley, 2016), ScriptaGrad provides a practical foundation for analysing student work beyond final-answer correctness. However, reliable recognition and error detection do not necessarily establish the causes of misunderstanding or determine the most appropriate instructional response. This limitation motivates the transition towards *EduHyperAgent*, which extends assessment with diagnostic dialogue, personalised interventions, and adaptive pedagogical orchestration.

The experience acquired with ScriptaGrad motivates a further change in the unit of design. A conventional assessment pipeline attempts to turn student work into a transcript, score or error category. The proposed *EduHyperAgent* instead treats the student's work as evolving, uncertain evidence for selecting a pedagogically appropriate next action. In this architecture, the teacher acts as a *pedagogical orchestrator*: setting curricular priorities, controlling permissible interventions, interpreting contested cases and deciding when a human consultation is necessary. Specialised services may handle handwriting recognition, spatial and formula parsing, symbolic verification, explanatory dialogue, oral follow-up, task generation and review scheduling. Their coordinated operation does not replace the teacher's professional authority.

The proposed system combines handwritten STEM evidence with optional spoken or typed explanations, task correctness, response latency, revision history and other carefully defined interaction indicators. These observations support a revisable *longitudinal learner state*, which is used to formulate further questions. For instance, both correct and incorrect written solutions may trigger an oral Socratic question: an incorrect answer can prompt a counterexample or a request for justification, while a correct answer can lead to a more demanding "why" or "under what conditions" question. Students who have demonstrated relevant prerequisites can be offered progressively deeper problems; students requiring foundational support receive smaller, structured steps and opportunities for successful independent practice. These are proposed policies to be evaluated, not claims that learners can or should be assigned fixed categories of ability.

We use *metacognitive adaptation* in a restricted engineering sense: the controller can evaluate uncertainty, monitor intervention outcomes, compare candidate policies offline, and propose changes for teacher-approved versions. It does *not* imply unrestricted self-rewriting software or autonomous changes to a deployed educational policy involving children. Nor does the framework infer temperament types (e.g., sanguine or phlegmatic), personality, or psychological condition from handwriting appearance. Such graphological inference lacks adequate validity for consequential educational decisions. Handwriting features are used to assess legibility, segmentation difficulty and recognition uncertainty, not to classify a child's character.

A central proposed mechanism is (GRPO), introduced in DeepSeekMath as a memory-efficient variant of policy optimisation using relative rewards for multiple candidate outputs (Shao et al., 2024). In our educational extension, a recognised error is associated not with one predetermined response but with a *group* of admissible prompts, hints, checks and escalation actions.

These candidates are scored against explicit criteria such as pedagogical suitability, factual validity, preservation of learner autonomy, accessibility, teacher-review burden and computational cost. Their group-relative evaluations may train a bounded routing policy offline. The weights and reward criteria must be calibrated and validated against expert teacher judgements; they have *not* been estimated from the historical school comparison, and no deployment of a trained GRPO controller is claimed here. Spline-based or coconvex approximation methods may additionally be investigated for representing temporal learning trajectories under appropriate shape constraints; whether they improve policy decisions is an empirical question rather than an established finding.

1.5 Research Objectives and Questions

The main objective of this study is to establish an evidence-informed conceptual and mathematical foundation for EduHyperAgent, a teacher-governed multi-agent framework for adaptive pedagogical decision-making.

Drawing on retrospective experience with ScriptaGrad, the study pursues three objectives:

1. To identify the capabilities and limitations of AI-assisted handwritten STEM assessment as a source of reliable pedagogical evidence.
2. To formulate a multimodal learner model integrating written solutions, optional oral diagnosis, and spline-based approximation of longitudinal learning trajectories.
3. To develop a constrained algorithm for selecting pedagogical strategies (PS-I or I-PS) and coordinating specialised AI agents, and to specify criteria for its future validation.

These objectives inform three research questions:

RQ1. What insights into the potential and limitations of AI-supported assessment of handwritten work does the ScriptaGrad retrospective offer for the design of the EduHyperAgent multi-agent platform?

RQ2. How can multimodal assessment evidence and spline-based learning trajectories support the mathematical representation of learner progress and diagnostic uncertainty?

RQ3. How can a teacher-governed decision algorithm use learner-state estimates to select adaptive instructional strategies and coordinate specialised pedagogical agents?

2. Background, Related Work and Theory-Building Synthesis

2.1 Agentic AI: Research Approaches and Applications.

The concept of an AI agent refers to a system capable of perceiving its environment, making decisions, and performing actions to achieve defined objectives, rather than merely generating conversational responses. Classical AI theory describes agents through their interactions with an environment, while multi-agent systems extend this framework through communication, coordination, and distributed decision-making (Russell and Norvig, 2021; Wooldridge, 2009).

Recent advances in large language models (LLMs) have expanded these capabilities by integrating reasoning processes, external tools, memory, and iterative planning. Unlike conventional LLMs that primarily generate responses to prompts, agentic systems can select actions, interact with specialised tools, evaluate intermediate results, and adapt subsequent decisions. However, advanced reasoning capabilities alone do not constitute agency; a defining characteristic is the ability to coordinate actions and respond to feedback within a defined operational environment.

Yao et al. (2023) introduce intertwined reasoning and environment actions (ReAct), while Schick et al. (2023) examine learning when and how to call external APIs. Such systems differ from merely prompting a model to write a lesson plan: they may invoke a symbolic solver, retrieve a rubric, detect a transcription ambiguity or trigger a teacher review. The multi-agent framework AutoGen illustrates coordination among configurable agents, people and tools (Wu et al., 2024). Yet decomposing one problem into many model roles can amplify propagation of errors, accumulate costs and obscure responsibility. Consequently, an educational "hyperagent" requires an explicit architecture and measurable advantages over a well-engineered single-agent or fixed-workflow baseline; the prefix *hyper* is a proposed orchestration layer, not a claim of superhuman or unrestricted autonomy.

A bounded educational agent episode is represented by the interaction history

$$h_t = (o_0, a_0, o_1, \ldots, a_{t-1}, o_t), \tag{2}$$

where o_t denotes an observation and a_t an authorised instructional action, including the selection of its executing agent.

The decision policy is defined as

$$a_t \sim \pi_\theta \left(\cdot \mid h_t, g, \mathscr{T}_t, \mathscr{D}_t \right), \tag{3}$$

where g is a teacher-approved instructional goal, $\mathscr{T}_t$ is the registry of authorised tools, and $\mathscr{D}_t$ represents the discovery state of candidate AI models and specialised micro-agents.

Following the selected action, the environment produces a new observation:

$$o_{t+1} \sim P(\cdot \mid h_t, a_t). \tag{4}$$

Unlike fixed-tool architectures, EduHyperAgent proposes governed dynamic agent discovery. New models and micro-agents may be identified and evaluated, but they enter the authorised tool registry only after technical validation, pedagogical assessment, and approval.

The architecture therefore combines adaptive tool selection with explicit constraints on privacy, instructional safety, learner autonomy, and teacher oversight.

Recent advances in agentic AI have extended language models from generating responses towards planning, tool use, and iterative decision-making. For example, *The AI Scientist* integrates hypothesis generation, programming, experimentation, and scientific reporting (Lu et al., 2024), while *OpenScholar* combines information retrieval, evidence-based synthesis, and iterative refinement (Asai et al., 2026). These systems demonstrate the potential of specialised tools and coordinated workflows. However, successful task execution does not necessarily imply reliable educational outcomes.

Intelligent tutoring systems have long integrated learner models, domain knowledge, and instructional strategies (VanLehn, 2011). Recent reviews of AI-supported tutoring emphasise pedagogical design, adaptive interaction, and human oversight (Létourneau et al., 2025; Liu et al., 2026; Li and Abdul Razak, 2026). Nevertheless, evidence that multi-agent architectures consistently outperform simpler tutoring approaches remains limited. In particular, immediate task correctness must be distinguished from conceptual understanding, learner autonomy, and delayed independent performance.

Building on these foundations, EduHyperAgent proposes a teacher-governed architecture that coordinates specialised components according to pedagogical needs. ScriptaGrad provides handwritten STEM mathematical evidence verification tools check recognised expressions; diagnostic agents support oral questioning; and instructional agents generate personalised learning tasks. A central orchestration policy selects appropriate interventions while preserving teacher control and student independence.

Unlike architectures restricted to a fixed collection of tools, EduHyperAgent also proposes governed dynamic agent discovery: a mechanism for identifying new AI models and specialised micro-agents, evaluating their technical and pedagogical suitability, and integrating approved components into the available tool registry. Discovery does not imply automatic deployment; new components require validation and authorised registration.

The proposed framework therefore separates four dimensions of evaluation: *technical task accuracy*, *pedagogical appropriateness*, *teacher oversight*, and *independent learning outcomes*. Table 1 summarises the relevant research foundations and their implications for EduHyperAgent.

Table 1. Research foundations and validation requirements for EduHyperAgent.

Research tradition	Relevant mechanism	Required validation
Scientific and tool-using agents	Planning, delegation, verification	Technical accuracy and traceability
Intelligent tutoring systems	Learner modelling and adaptive support	Conceptual understanding and transfer
Multi-agent orchestration	Specialised tools and dynamic discovery	Pedagogical suitability and reliability
Teacher–AI collaboration	Human oversight and intervention control	Learner autonomy and independent mastery

The central research question is therefore not whether more agents can complete more tasks, but whether their coordinated use improves pedagogical decisions and supports independently demonstrated learning. These benefits remain subject to prospective experimental validation.

2.2 Critical comparison of recent agentic education reviews

Reviews published in 2026 substantially refine the scope of this field. A systematic synthesis of 33 empirical studies of generative agents in education identified cognitive–epistemic, self-regulatory, affective and human–AI interaction roles (Wang et al., 2026). This functional typology is useful, but role proliferation is not proof of causal benefit. A separate scoping review

mapped 474 studies published between January 2020 and May 2026, analysing research designs and dimensions of agentic capability (Wang et al., 2026). Its breadth sharpens an important distinction: architectural agency and educational effectiveness are different evidential questions. A European Commission systematic review of causal evidence on intelligent tutoring systems found positive but non-uniform learning effects, with more heterogeneous motivation and socio-emotional evidence (Theis et al., 2026). The review concerns ITS broadly and should not be interpreted as a meta-estimate for LLM multi-agent systems.

The recently published integrative review by Li and Abdul Razak (2026) examined personalisation, co-design and ethical mediation in higher education, whereas Lin et al. (2026) and (Zalizko, 2026b) mapped generative-AI practices and research outcomes more broadly. These higher-education findings cannot be assumed to transfer without qualification to primary pupils, emerging literacy, differently developed self-regulation or mathematics notation across grades 1–11. In programming education, a 2026 evidence review identified over-reliance and assessment integrity as explicit dimensions alongside learning gains (Kumar et al, 2026). Accordingly, we operationalise agency as tool authority, traceability, human override and preservation of independent pupil work, and stratify evaluation by grade band rather than pooling heterogeneous ages without justification.

Table 2. Research coverage and remaining gaps in recent AI education reviews.

Review	Agents	K–12	Outcomes	Governance
Liu et al. (2026)	+	○	○	○
Agentic review (2026)	+	○	○	○
Theis et al. (2026)	–	+	+	○
Li & Razak (2026)	+	–	○	+
Lin et al. (2026)	○	–	○	○
EduHyperAgent (proposed)	+	+	+	+

Note. + = central research focus; ○ = addressed but not central; – = outside the main scope. Coding is an interpretive synthesis, not a quantitative quality or effect assessment. The EduHyperAgent row denotes intended research coverage, not validated findings.

Table 2 highlights the fragmentation of current research across agentic architectures, K–12 education, learning outcomes and teacher governance. EduHyperAgent aims to connect these dimensions through uncertainty-aware evidence processing and pedagogically constrained multi-agent routing.

2.3 Multimodal Assessment and Pedagogical Evidence

Handwritten STEM solutions provide valuable evidence of students' reasoning, but their interpretation depends on image quality, mathematical notation, diagrams, and intermediate solution steps. Multimodal learning analytics extends conventional activity logs to richer forms of educational evidence (Blikstein and Worsley, 2016; Mu et al., 2020). However, accurate recognition does not necessarily imply reliable conceptual diagnosis.

Recent studies have investigated image-based mathematical assessment, AI-assisted grading, and spline-guided document segmentation (Kapur, 2014; Baral et al., 2023; Kortemeyer et al., 2025; Zalizko, 2026b). These approaches motivate ScriptaGrad as a specialised component for analysing handwritten STEM work. EduHyperAgent extends this foundation by combining written evidence with oral diagnostic questions and longitudinal learner modelling, while preserving uncertainty and teacher review.

2.4 Adaptive Learning, Scaffolding, and Teacher Oversight

Research on scaffolding and formative feedback emphasises support that responds to students' current understanding rather than providing complete solutions by default (Wood et al., 1976; Hattie and Timperley, 2007; VanLehn, 2011). The ICAP framework further highlights the importance of meaningful cognitive engagement (Chi and Wylie, 2014).

Studies of Productive Failure distinguish problem solving followed by instruction (PS-I) from instruction followed by problem solving (I-PS), showing that instructional sequencing can influence conceptual understanding (Sinha and Kapur, 2021). Recent research on pedagogically guided LLM tutoring demonstrates the potential of structured instructional strategies, while also identifying risks of excessive AI assistance (Puech et al., 2025; Bastani et al., 2025).

Learner modelling offers a basis for adapting these strategies over time. Knowledge tracing and open learner models represent learning progress while supporting the interpretation and correction of learner-state estimates (Corbett and Anderson, 1994; Bull and Kay, 2016). Human–AI collaboration further requires transparent decisions and meaningful teacher control (Holstein et al., 2019; Molenaar, 2022).

Building on these principles, EduHyperAgent proposes using written and oral evidence to select individualised instructional pathways while preserving students' independent reasoning.

2.5 Multi-Agent Orchestration and Pedagogical Optimisation

Tool-using AI agents demonstrate how complex tasks can be decomposed into specialised operations involving retrieval, verification, and iterative decision-making (Yao et al., 2023; Schick et al., 2023; Wu et al., 2024). However, coordinating multiple models does not necessarily improve accuracy or educational effectiveness.

EduHyperAgent proposes a teacher-governed orchestration framework that selects specialised AI tools and instructional interventions according to pedagogical objectives, diagnostic evidence, and operational constraints. The architecture also considers controlled discovery and validation of additional specialised agents, rather than relying exclusively on a fixed tool registry.

For policy optimisation, GRPO offers a potential method for comparing candidate actions through relative rewards (Shao et al., 2024; Guo et al., 2025). Unlike mathematically verifiable answers, pedagogical quality requires multidimensional evaluation, including instructional appropriateness, learner autonomy, technical reliability, and teacher oversight. Consequently, improvements in optimisation rewards cannot be interpreted directly as learning gains.

The proposed framework therefore separates technical performance from educational effectiveness. Its central hypothesis is that coordinated, evidence-based selection of specialised agents can support more appropriate instructional decisions than fixed AI assistance. This hypothesis requires prospective experimental validation.

2.6 Defining the research gap

Across these literatures, recognition, assessment, tutoring and learner modelling are usually evaluated with different endpoints. In particular, good handwritten transcription does not show that an ensuing hint preserves reasoning; a successful tutoring interaction does not establish delayed transfer; and low per-call compute does not establish lower teacher workload. The testable gap is a *joined causal and operational account*: which uncertain piece of learner evidence justified which opportunity, how that support affected independent learning, and at what burden to the teacher? Table 3 specifies the empirical consequence of each research strand.

Table 3. Research carried out in the past and requirements for a design that can be disproven.

Research strand	Design requirement	Disconfirming evidence
Handwritten STEM analysis	Keep alternative transcriptions and assessment uncertainty	No gain over direct VLM and no useful review–risk trade-off
Scaffolding/PS-I	Bounded support at episode boundaries	Adaptation reduces delayed independent performance
Open learner modelling	Inspectable, correctable state and evidence provenance	Corrections fail to change downstream decisions
Affordances/action traces	Record available, admissible, exposed and enacted support	Independent auditors gain no interpretive advantage
Efficient routing	Choose least costly admissible tool	Savings disappear after training, escalations and teacher corrections

3. Methodological Foundations of EduHyperAgent

This section establishes the methodological foundations for developing EduHyperAgent. It specifies how complementary learning evidence can inform learner-state modelling and adaptive instructional decisions. Spline-based trajectory approximation and constrained pedagogical strategy selection provide the mathematical basis for the proposed framework.

3.1 Integrating Written, Oral, and Longitudinal Learning Evidence

Learning analytics draws on multiple educational data sources, yet more data do not necessarily provide a better understanding of learning processes (Blikstein and Worsley, 2016; Mu et al., 2020). The methodological challenge is to distinguish observable performance from underlying knowledge and reasoning.

In STEM education, handwritten solutions reveal mathematical procedures and possible errors but may not distinguish conceptual misconceptions from procedural mistakes. Oral diagnostic questions provide complementary evidence, while longitudinal records help characterise individual learning progress.

The proposed EduHyperAgent methodology integrates these sources through ScriptaGrad and specialised pedagogical agents. Learning evidence is represented not as isolated scores but as a sequence of observations supporting diagnosis, trajectory modelling, and instructional strategy selection.

Table 4 summarises the proposed evidence structure.

Table 4. Learning evidence and pedagogical functions in the proposed EduHyperAgent framework.

Evidence source	Analytical function	Pedagogical application
Handwritten STEM work	Recognition, error detection, mathematical verification	Identification of potential misconceptions
Oral explanations	Diagnostic questioning, conceptual clarification	Selection of instructional support
Learning history	Progress monitoring, trajectory approximation	Personalised task sequencing
Independent reassessment	Verification of understanding and knowledge transfer	Learner-state update and teacher review

These sources are complementary but uncertain. Recognition accuracy does not establish conceptual understanding, and correct explanations do not necessarily demonstrate independent mastery. Consequently, the proposed methodology preserves diagnostic uncertainty, incorporates teacher review, and requires independent reassessment.

3.2 EduHyperAgent: Formal Pedagogical Architecture

EduHyperAgent is conceived as a teacher-governed multi-agent architecture in which pedagogical decision-making is separated from task execution. The framework not only routes learner requests to currently registered specialist agents, but also supports the discovery and local evaluation of new educational AI tools for narrowly defined instructional functions in Figure 5.

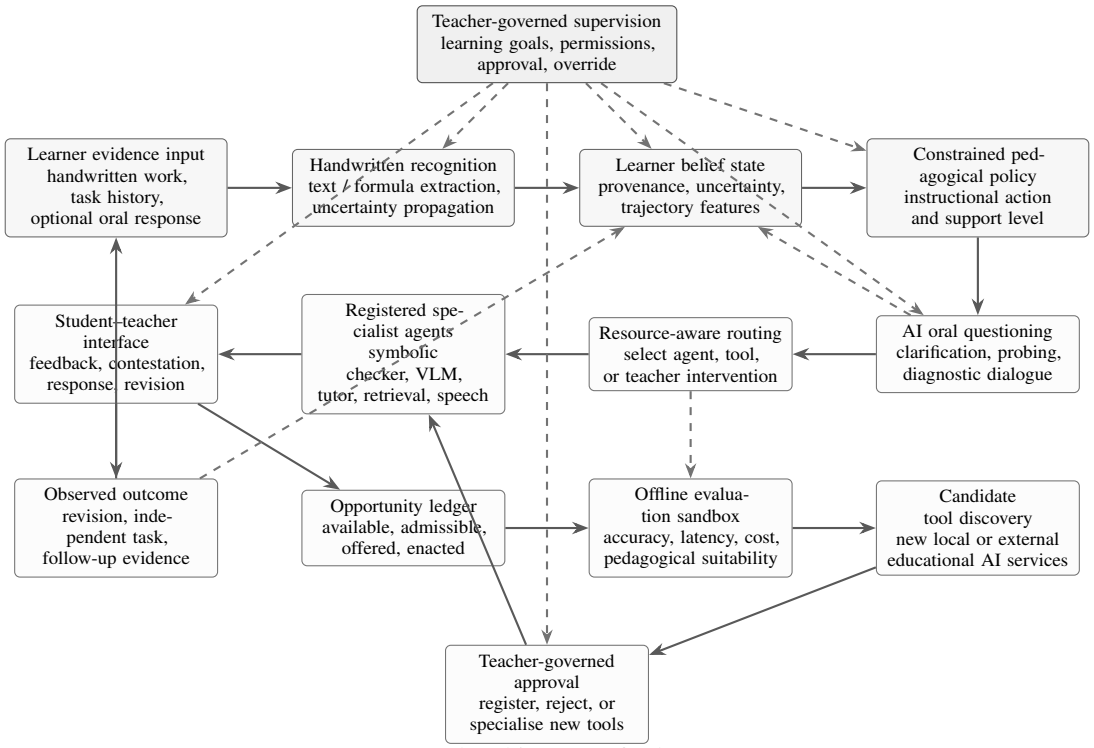


Figure 5. Formal architecture of EduHyperAgent.

The upper layer represents teacher-governed supervision over learner-state estimation, pedagogical policy, AI-generated oral questioning, and tool admissibility. The core loop integrates handwritten recognition, multimodal learner-state updating, constrained pedagogical action selection, specialist-agent routing, and subsequent learner response. The lower layer supports the discovery, offline evaluation, and controlled registration of new educational AI tools. Dashed arrows indicate supervisory, corrective, and approval relations rather than ordinary execution flow. The architecture therefore distinguishes between two levels of adaptation. At the instructional level, the system selects pedagogical actions for a given learner state. At the infrastructural level, it searches for and evaluates specialised AI services that may better solve narrowly defined educational

tasks, such as handwritten formula recognition, symbolic verification, oral clarification, hint generation, or adaptive practice selection.

This distinction is important because the set of available educational AI tools should not be treated as fixed. A new tool may improve performance for one specific subtask while remaining unsuitable for broader pedagogical use. Accordingly, candidate tools are first assessed offline with respect to accuracy, latency, computational cost, teacher-review burden, and pedagogical suitability before they are registered as admissible agents within EduHyperAgent.

The resulting framework enables local optimisation of task execution without surrendering pedagogical control. Instructional policy remains teacher-governed, whereas specialised AI agents act as bounded executors of authorised actions.

3.3 Evidence-Based Adaptive Pedagogical Decision-Making

Building on the learning analytics cycle and adaptive instruction principles, the proposed methodology links learning evidence to pedagogical decisions (Hattie and Timperley, 2007; VanLehn, 2011).

For learner i at instructional episode t, multimodal evidence is represented as

$$\mathscr{E}_{it} = \{W_{it}, O_{it}, H_{it}\}, \tag{5}$$

where W_{it} denotes written work, O_{it} optional oral diagnostic responses, and H_{it} longitudinal learning history.

The learner's progress is represented by a trajectory approximated using spline functions fitted to longitudinal assessment observations. This modelling approach is intended to capture individual changes in performance while reducing the influence of isolated measurement fluctuations.

Learner-state estimates derived from multimodal evidence and trajectory modelling inform a constrained pedagogical decision algorithm. The algorithm selects instructional strategies and specialised agents according to estimated learning needs, diagnostic uncertainty, and teacher-approved objectives.

Admissible pedagogical actions belong to $\mathscr{A}^{\mathrm{adm}}(b_{it})$ and satisfy curricular, privacy, accessibility, and pedagogical safety constraints. Possible actions include further diagnosis, personalised practice, teacher review, and selection between problem solving followed by instruction (PS-I) and instruction followed by problem solving (I-PS) (Sinha and Kapur, 2021; Puech et al., 2025).

The resulting written–oral–written cycle provides a conceptual mechanism for updating learner-state estimates and adapting subsequent interventions. Agent selection and instructional actions remain subject to teacher oversight.

This framework represents a proposed methodological foundation rather than an experimentally validated adaptive learning system. Its development is informed by retrospective ScriptaGrad experience, examined in the following section.

3.4 A revisable partially observed learner state

Define an underlying learner state $Z_t = (K_t, M_t, R_t, U_t, G_t)$ comprising skill knowledge, possible misconceptions, reasoning/self-revision characteristics, uncertainty and current instructional goals. The learner state is not observed directly. Based on observations $O_{1:t}$ and prior interventions $A_{1:t-1}$, the system maintains a belief distribution

$$b_t(z) = p(Z_t = z \mid O_{1:t}, A_{1:t-1}). \tag{6}$$

A generative update has the form

$$b_{t+1}(z') \propto p(O_{t+1} \mid z', A_t) \int p(z' \mid z, A_t) b_t(z)\, dz. \tag{7}$$

This is a specification of the inference problem, not a claim that a fully identified Bayesian state-space model is already deployed. An operational approximation may use interpretable factors, calibration layers and a versioned evidence ledger. Every inferred diagnosis must retain the underlying handwriting excerpt, recognition alternatives, rubric version, model version, temporal scope and override history. Student requests for review and teacher corrections induce a new state estimate rather than preserving erroneous classifications.

3.5 Temporal indicators and instructional episode boundaries

Let $x_j(t)$ be task-normalised observations of independently completed substeps, productive revisions or justified strategies. A regularised spline $S_j(t)$ is fitted only with observations available before a decision at t. A proposed readiness summary is

$$q_t = \mathrm{sigmoid}\left(\beta_0 + \sum_j \{\beta_j S_j(t) + \gamma_j S_j'(t) + \delta_j S_j''(t)\}\right). \tag{8}$$

This score is a *decision indicator*, not a validated probability of a student's aptitude or willingness to struggle. For sparse or irregular measurements, derivatives can amplify noise; future evaluation should compare the spline-based indicators against last-observation, exponentially weighted summary and no-spline baselines. Shape-preserving constraints apply only to indicators with justified shape restrictions: actual learning trajectories may rise, fall or plateau. Missing optional speech does not lower readiness by definition.

A transparent mode selector applies at *episode boundaries*, not individual generated tokens:

$$\pi_t = \begin{cases} \text{PS-I}, & q_t \geq \tau_H \text{and prerequisites are met}, \\ \text{I-PS}, & q_t \leq \tau_L, \\ \pi_{t-1}, & \tau_L < q_t < \tau_H \text{ and a previous mode exists}, \\ \pi_{\text{default}}, & \text{otherwise}, \end{cases} \qquad 0 \leq \tau_L < \tau_H \leq 1. \tag{9}$$

PS-I always includes a subsequent consolidation stage and I-PS includes meaningful independent practice. Thresholds must be estimated on development data and frozen for confirmatory evaluation. The initial or fallback mode π_{default} is teacher-approved. High diagnostic uncertainty triggers clarification or teacher review before the selected mode is enacted.

3.6 Available, admissible, offered and enacted action records

Let $\mathscr{A}_t^{\text{avail}}$ denote technically executable actions, $\mathscr{A}_t^{\text{adm}} \subseteq \mathscr{A}_t^{\text{avail}}$ actions allowed under curriculum, safety, privacy and episode constraints, and $\mathscr{A}_t^{\text{offer}} \subseteq \mathscr{A}_t^{\text{adm}}$ those actually exposed to the learner. An enacted action a_t can also be independent continuation or teacher override. The event ledger is

$$\Lambda_t = (b_t, \mathscr{A}_t^{\text{avail}}, \mathscr{A}_t^{\text{adm}}, \mathscr{A}_t^{\text{offer}}, a_t, O_{t+1}, h_t), \tag{10}$$

where h_t records rationale, model and rule versions, provenance and human override. An action omitted from the offer set cannot be treated as a learner-rejected opportunity. Conversely, a visible option that is not clicked does not prove it was understood or perceived; additional observation is needed. This record is a proposed extension of learning-analytics event semantics rather than a causal account of learning by itself.

Within the selected instructional mode, assistance should progress from independent attempts and diagnostic prompts to targeted hints or worked explanations, according to evidence from learners and teacher-approved rules. The aim is to provide sufficient support while preserving opportunities for independent reasoning. Decisions and teacher overrides should be recorded for subsequent evaluation.

4. Results: Retrospective Evidence and Methodological Foundations of EduHyperAgent

Retrospective evidence provides a description of AI-supported educational feedback, whereas the EduHyperAgent components presented below constitute a method that requires prospective validation. Accordingly, no causal learning effect is attributed to the proposed architecture.

The retrospective class-level comparisons are reported first; the proposed spline estimator, pedagogical selector, and resource-routing scheme are then specified as separate methodological contributions. The aggregated retrospective data cannot validate individual trajectories or adaptive policy effects.

4.1 Retrospective Dataset and Indicator-Level Contrasts

The article draws on educational feedback data collected at the International Innovative School between 2019 and 2024. The dataset comprises eleven indicators of conventional and AI-supported feedback across eight class groups (grades 5, 6, 7, 8A, 8B, 9, 10 and 11), with twelve observations recorded per indicator, group and feedback condition.

The retrospective analysis uses class-group means to identify descriptive differences in feedback characteristics and educational performance indicators. These findings provide an empirical foundation for identifying methodological requirements for AI-supported assessment and adaptive pedagogical decision-making.

Building on this retrospective evidence, EduHyperAgent is proposed as a teacher-governed framework integrating handwritten STEM assessment, optional oral diagnosis, longitudinal learner modelling, and specialised pedagogical agents. Future research will extend the empirical evaluation to individual longitudinal learning trajectories, diagnostic reliability, adaptive instructional strategies, and independently assessed learning outcomes.

Table 5. Descriptive retrospective comparison across eight class groups. Differences are AI-supported minus conventional group means; units and construct directions are not interchangeable.

Indicator (unit)	Conventional	AI-supported	Difference
Feedback timeliness (score)	6.26	8.85	+2.59
Feedback readability (score)	8.26	7.23	−1.03
Additional explanations (count)	9.03	29.21	+20.18
Visualisation quality (score)	9.10	8.03	−1.07
Confusing explanations (count)	2.23	2.60	+0.38
Motivation (reported %)	66.98	77.19	+10.21
Error correction (reported %)	38.49	54.65	+16.16
Summary tests (score)	7.81	9.09	+1.28
Student engagement (score)	9.13	6.73	−2.40
Student satisfaction (score)	8.13	7.67	−0.46
Feedback individualisation (score)	8.02	8.73	+0.71

For indicator k, group g, and condition c, let $\bar{y}_{kgc}$ denote the arithmetic mean of the available numerical entries. The class-level contrast is $\Delta_{kg} = \bar{y}_{kg,\mathrm{AI}} - \bar{y}_{kg,\mathrm{conventional}}$. Table 5 reports the unweighted means across the eight groups. These values retain their original units and must not be combined without a validated scoring and weighting scheme.

The higher reported timeliness, frequency of explanations, error correction, and summary-test scores occurred alongside lower reported readability, visualisation quality, engagement, and satisfaction. The indicator "confusing explanations" also increased, for which a positive numerical contrast is not a desirable outcome. All eight groups showed higher mean correction percentages and summary-test scores in the AI-supported condition; all eight showed lower mean engagement ratings.

Standardised differences between AI-supported and conventional feedback across eight class groups. Indicator codes: TIM = feedback timeliness; READ = feedback readability; EXPL = additional explanations; VIS = visualisation quality; CONF = confusing explanations; MOT = student motivation; CORR = error correction; TEST = summary test performance; ENG = student engagement; SAT = student satisfaction; PERS = feedback individualisation. Positive differences indicate higher recorded values, not necessarily better educational outcomes.

4.2 Exploratory Multivariate Structure of Feedback Indicators

A principal component analysis (PCA) was performed on the 8×11 matrix of *AI-supported class-group means*. To avoid giving larger numerical units disproportionate influence, each indicator was centred and divided by its sample standard deviation before singular-value decomposition. The first three components explained 40.30%, 22.58%, and 14.27% of the standardised class-level variance, respectively (77.15% cumulatively); the first two explained 62.88%. The two-dimensional PCA biplot is presented at the end of the Results section as a synthesis of the retrospective findings. An interactive three-dimensional rendering is provided as supplementary material for exploratory inspection. The plotted directions are associations among *group-level indicators*, not latent cognitive mechanisms or causal links. PCA is particularly unstable here because there are only eight groups and eleven indicators: after centring, the data matrix has rank at most seven. Component directions and apparent group separation must therefore be treated as exploratory rather than generalisable.

4.3 Empirical Implications for ScriptaGrad and EduHyperAgent

The observed contrasts delineate a concrete design problem. AI-supported feedback was associated with higher recorded explanation frequency and error-correction percentages, but not with uniformly higher ratings of feedback clarity, engagement, or satisfaction. Thus, producing more responses and recording more corrected errors cannot substitute for checking the intelligibility, appropriateness, and educational consequences of feedback. The retrospective database contains feedback indicators and test ratings, not independent validation of handwritten recognition accuracy, reliable misconception diagnosis, or the effect of adaptive agent selection.

These limitations motivate three distinct functional requirements. First, ScriptaGrad should retain the original handwritten solution, transcription alternatives, mathematical-verification status, and teacher corrections as auditable evidence rather than reduce a response immediately to a single score. Second, EduHyperAgent should combine written evidence with optional oral clarification and comparable longitudinal observations while representing missingness and uncertainty explicitly. Third, pedagogical strategy selection should depend on the estimated learner state and teacher-approved goals, with an independent follow-up task used to update that state. These requirements are derived design implications, not results of a completed

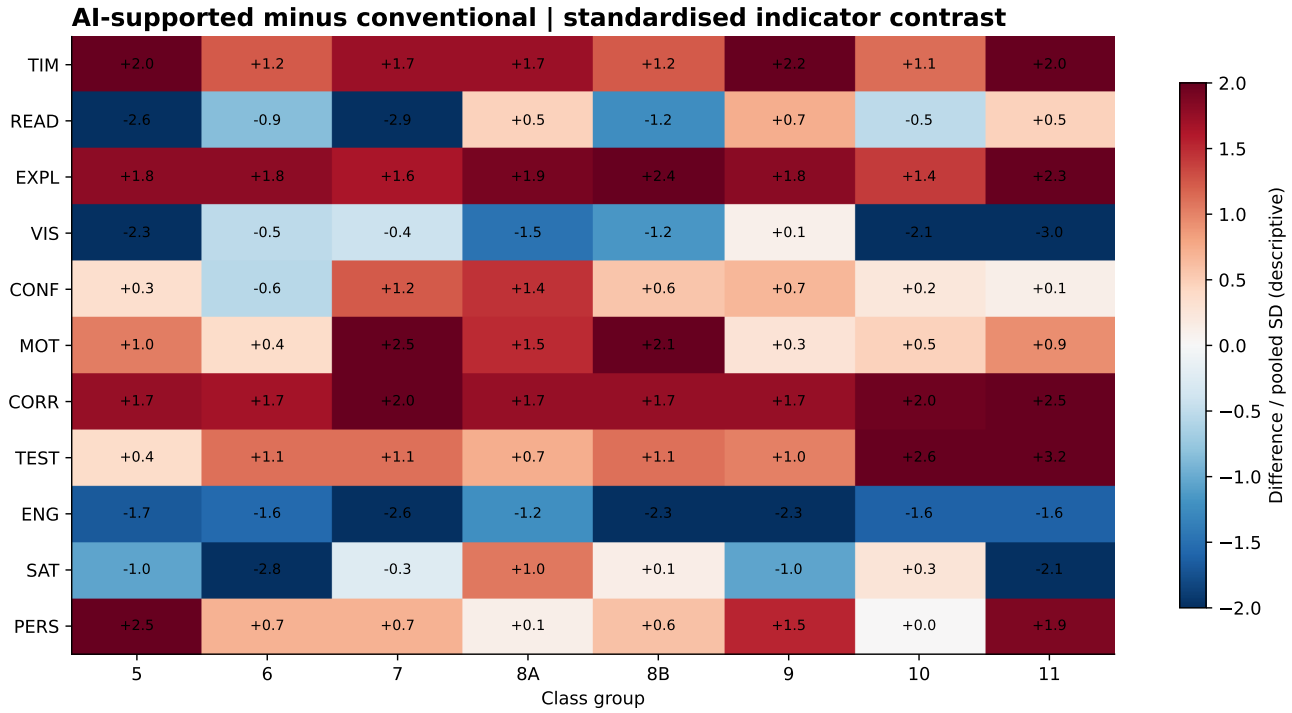


Figure 6. Standardised descriptive contrasts in the eleven indicators for eight class groups. Cell values are condition differences divided by the pooled standard deviation of the sixteen group-condition means for the corresponding indicator; positive values indicate higher recorded values, not necessarily better outcomes.

EduHyperAgent intervention.

The correspondence with the research questions is explicit: the indicator contrasts and exploratory PCA address RQ1 by characterising the opportunities and limitations of historical AI-supported feedback; the evidence representation and proposed trajectory model address RQ2; and the constrained instructional policy and specialist routing address RQ3. None of these results establishes a prospective learning gain from EduHyperAgent.

4.4 Spline-Based Learner Trajectories and Pedagogical Strategy Selection

For learner i, let $\{(\tau_{ij}, y_{ij}, w_{ij})\}_{j=1}^{n_i}$ denote longitudinal observations, where τ_{ij} is assessment time, y_{ij} is a comparable performance indicator, and w_{ij} represents observation reliability. The proposed trajectory is estimated by a cubic smoothing spline:

$$\widehat{S}_i = \arg\min_{S \in \mathscr{S}_3} \left\{ \sum_{j=1}^{n_i} w_{ij} \left[y_{ij} - S(\tau_{ij}) \right]^2 + \lambda_i \int_{\tau_{i1}}^{\tau_{in_i}} \left[S''(u) \right]^2 du \right\}. \tag{11}$$

The local learner-trajectory features are

$$\mathbf{z}_i(t) = \left[\widehat{S}_i(t), \widehat{S}_i'(t), \widehat{S}_i''(t) \right]^\top. \tag{12}$$

These quantities describe estimated performance level, local change, and curvature of the fitted trajectory; they are not interpreted as direct measures of cognitive development or causal learning effects. Multimodal evidence and trajectory information define the proposed learner state

$$b_{it}(\cdot) = \Phi(W_{it}, O_{it}, H_{it}, \mathbf{z}_i(t), u_{it}). \tag{13}$$

where W_{it} is written evidence, O_{it} optional oral evidence, H_{it} learning history, and u_{it} diagnostic uncertainty. Here, Φ returns a belief distribution over the learner state, consistent with Equation (6); the learner index was omitted there for readability. This mapping is a proposed inference mechanism, not an estimator fitted in the present study.

Pedagogical strategy selection is then formulated as

$$a_{it}^{*} = \arg \max_{a \in \mathscr{A}_{\text{adm}}(b_{it})} \widehat{U}_{\text{ped}}(a \mid b_{it}), \tag{14}$$

Equation (14) specifies the general decision objective. The threshold rule in Equation (9) is a proposed interpretable approximation for selecting the instructional mode.

4.5 Resource-Aware Routing and Offline Optimisation

After the pedagogical action is selected, an action–resource pair $\xi = (a,m)$ associates the action with an authorised executor m. Hard constraints first define the feasible set $\mathscr{F}(b_{it})$. Within that set,

$$\xi_{it}^{*} = \arg \min_{\substack{\xi=(a,m)\in\mathscr{F}(b_{it}) \\ a=a_{it}^{*}}} J(\xi \mid b_{it}). \tag{15}$$

with

$$J(\xi \mid b_{it}) = \lambda_C C(\xi) + \lambda_L L(\xi) + \lambda_T T(\xi) - \widehat{U}_P(\xi \mid b_{it}). \tag{16}$$

Here, C, L, and T denote measured computational cost, latency, and expected teacher-review burden, while $\widehat{U}_P$ denotes estimated pedagogical suitability. Safety, privacy, accessibility, minimum validity, and teacher-authorisation constraints are applied before cost optimisation; a cheaper action cannot compensate for an inadmissible one. The pedagogical action a_{it}^{*} is held fixed during executor selection. If no admissible executor is available, the decision is referred to the teacher. The cost and suitability terms require predefined scaling and independently calibrated weights.

Independent reassessment without AI assistance should evaluate mastery, justification, and transfer using predefined rubrics. Prospective validation should examine rater agreement and score comparability across tasks and learner groups before combining ratings into a composite measure.

4.6 Synthesis of Findings and Spline-Informed GRPO Optimisation

The findings address three complementary research objectives and establish the methodological basis for spline-informed, teacher-governed optimisation of pedagogical decisions. The retrospective results provide empirical motivation for the framework, whereas its proposed optimisation mechanisms remain subject to prospective validation. Figure 7 summarises the exploratory covariance structure of the historical feedback indicators.

For RQ1, the eleven-indicator retrospective comparison (Table 5) reveals both opportunities and limitations of historical AI-supported feedback. Recorded error correction and summary test scores were higher, whereas engagement and several clarity-related indicators were lower. The exploratory principal component analysis (PCA) characterises variation among feedback indicators but does not establish causal effects or instructional superiority.

For RQ2, the proposed learner-state estimator integrates written evidence, optional oral diagnostics, and longitudinal learning history (Equation (13)). The penalised smoothing spline in Equation (11) provides an estimated learning trajectory, while the features in Equation (12) describe its current level, local rate of change, and curvature. These features are intended to distinguish, for example, sustained progress, stagnation, and potentially unstable performance. Such interpretations remain provisional and depend on the quality and density of individual longitudinal observations. The spline model has not been fitted to the retrospective class aggregates.

For RQ3, the proposed pedagogical policy (Equation (14)) uses the estimated learner state to select among teacher-authorised instructional actions, including problem solving before instruction (PS-I), instruction before problem solving (I-PS), diagnostic clarification, targeted practice, and teacher referral. Specialist-agent selection is additionally governed by the admissibility constraints and cost–suitability criteria in Equations (15)–(16).

A central prospective extension is the use of weighted GRPO to improve these decisions through comparative evaluation of candidate actions. Rather than assigning a uniform reward structure to every learner and instructional situation, EduHyperAgent proposes to condition reward weights on the spline-derived learner state. For example, limited progress may increase the relative importance of diagnostic clarification and instructional appropriateness, whereas stable progress may favour independent practice and knowledge transfer. These relationships represent proposed policy-design principles, not empirically established decision rules.

The optimisation mechanism can be expressed conceptually as a state-dependent reward: $R_i(\xi,t) = \sum_{k=1}^{K} \omega_k\big(\mathbf{z}_i(t)\big) r_k(\xi, b_{it})$. where $\xi = (a,m)$ denotes a candidate action–executor pair, and b_{it} is the learner-state belief distribution and $\mathbf{z}_i(t)$ contains

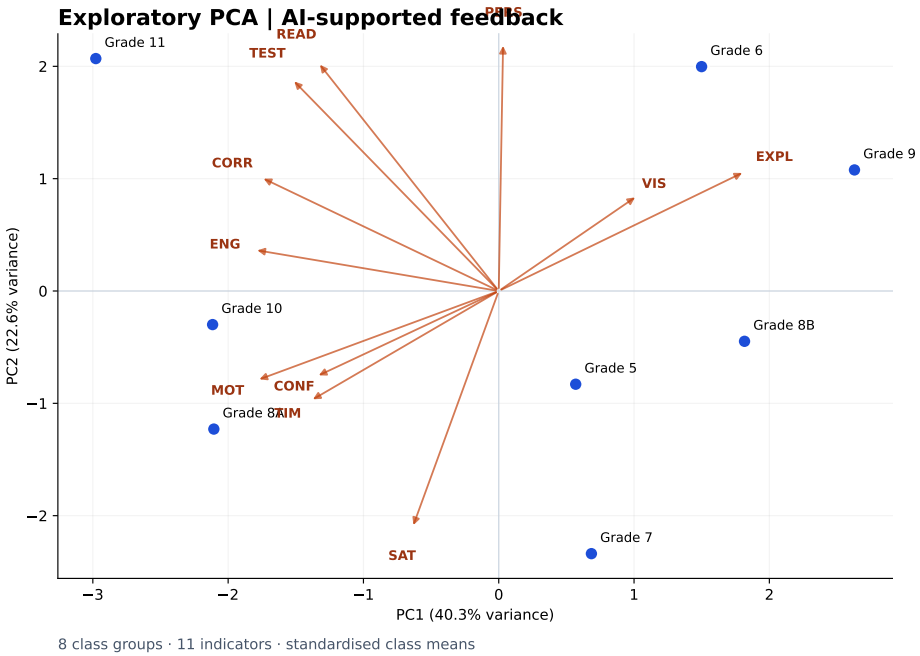


Figure 7. Exploratory two-dimensional PCA biplot of eleven standardised feedback indicators from eight class groups in the historical AI-supported condition (PC1: 40.30%; PC2: 22.58%). Points represent class groups and arrows represent indicator loadings. TIM: timeliness; READ: readability; EXPL: explanations; VIS: visualisation; CONF: confusing explanations; MOT: motivation; CORR: error correction; TEST: summary tests; ENG: engagement; SAT: satisfaction; PERS: individualisation. The analysis is descriptive and exploratory (eight groups, eleven variables), not causal or learner-level.

the spline-derived trajectory features. The term r_k represents a reward component, such as pedagogical suitability, diagnostic reliability, subsequent learning progress, computational efficiency, or teacher-review burden. The non-negative weights ω_k determine the relative importance of these components under the current learner state and are normalised to sum to one.

Within the proposed GRPO procedure, multiple admissible candidate decisions would be evaluated for comparable learner states. Their weighted rewards would provide group-relative advantages for updating the pedagogical and routing policy. Teacher-defined constraints would remain binding, and teacher assessments could contribute to the reward specification and offline evaluation. This separation is important: spline modelling estimates learning dynamics, reward weighting expresses context-dependent pedagogical priorities, and GRPO optimises the policy governing instructional and computational decisions.

The intended benefit is a more selective allocation of pedagogical interventions and computational resources, avoiding unnecessary specialist-agent invocations while preserving instructional appropriateness and teacher oversight. However, neither the state-dependent reward weights nor the GRPO policy have been trained or validated in the present study. Their effectiveness, stability, and resource implications require prospective learner-level data and controlled comparative evaluation.

5. Discussion

5.1 Interpretation of Retrospective Findings

The retrospective analysis of AI-supported educational feedback conducted between 2019 and 2024 revealed both positive tendencies and limitations of automated instructional support. Compared with conventional feedback, the AI-supported condition showed higher average values for feedback timeliness, error correction, summary-test performance, and feedback individualisation. In particular, the recorded error-correction indicator increased from 38.49% to 54.65%, while average summary-test scores increased from 7.81 to 9.09.

However, these improvements were accompanied by lower reported student engagement, feedback readability, and satisfaction. Consequently, the retrospective results do not support an assumption that increasing AI assistance automatically improves all dimensions of learning.

The exploratory principal component analysis further demonstrated that the feedback indicators contain substantial

multidimensional variation. The first two principal components accounted for 62.88% of standardised variance across the eight class groups. Given the small number of aggregated observations, this finding is descriptive rather than confirmatory.

Together, these results suggest that automated feedback should be evaluated not only by its speed or the frequency with which errors are corrected, but also by its pedagogical relevance, diagnostic reliability and impact on independent student learning.

5.2 Mathematical Contribution: Spline-Based Learning Analytics

The mathematical contribution lies in linking longitudinal learner modelling to instructional decisions. The spline formulation in Equation (11) combines reliability-weighted observations with regularisation, while Equation (12) represents estimated performance level, change, and curvature. These features describe assessment trajectories, not cognitive mechanisms or causal learning effects.

Optional shape-preserving or coconvex constraints may represent theoretically justified local patterns without imposing globally monotonic progress. Their usefulness must be evaluated against ordinary smoothing splines and simpler longitudinal models.

5.3 Adaptive Strategy Selection and Multi-Agent Orchestration

The second methodological contribution links learner-state estimation to pedagogical strategy selection. Equation (13) integrates written evidence, optional oral diagnosis, learning history, spline-derived features, and diagnostic uncertainty. Equations (9) and (14) describe how this representation informs teacher-authorised instructional decisions, including PS-I/I-PS selection, further diagnosis, and independent practice.

Pedagogical strategy selection precedes the choice of its executor. Equations (15) and (16) specify routing according to estimated pedagogical suitability, computational cost, latency, and teacher-review burden. Safety, accessibility, privacy, and minimum technical reliability remain hard constraints. New agents may undergo offline benchmarking before teacher-authorised registration and deployment.

The mixed retrospective feedback findings motivate more selective instructional support, but do not establish the benefits of adaptive routing. Weighted GRPO is proposed as a future mechanism for optimising pedagogical and resource-allocation decisions within these constraints. Its effects on instructional quality, resource consumption, and teacher workload require controlled comparative evaluation.

5.4 Methodological Contributions and Research Implications

The proposed methodology separates recognition, mathematical verification, pedagogical diagnosis, strategy selection, and task execution, supporting error tracing and teacher review. Spline-based modelling complements isolated scores with longitudinal indicators, while separating instructional decisions from resource routing reduces dependence on individual AI models. An agent registry enables controlled integration of specialised tools, with teacher oversight, evidence provenance, and independent reassessment embedded in the design.

These features support modularity and auditability; improvements in learning outcomes, diagnostic accuracy, or efficiency remain to be demonstrated. The mixed retrospective findings motivate investigating when and how these mechanisms provide educational value.

5.5 Limitations and Future Validation

The study combines retrospective descriptive analysis with conceptual and mathematical development. Aggregated class-group comparisons cannot establish causal effects, and PCA remains exploratory with only eight groups.

Validating spline-based trajectories requires individual longitudinal data, including evaluation of smoothing parameters, derivative stability, and optional curvature constraints. Pedagogical utility and routing parameters also require independent calibration.

Future studies should compare spline-informed and simpler learner models, fixed and adaptive PS-I/I-PS selection, and alternative routing strategies under matched instructional conditions. Primary outcomes should include independently assessed mastery and delayed transfer, with diagnostic reliability, teacher workload, computational efficiency, and learner autonomy evaluated separately.

6. Conclusion

This study develops an evidence-informed conceptual and mathematical foundation for EduHyperAgent. Retrospective analysis of AI-supported feedback from 2019 to 2024 revealed higher recorded timeliness, error correction, test performance,

and individualisation alongside lower engagement and selected feedback-quality ratings. These mixed findings motivate instructional support that considers both performance and the quality of learning evidence.

The central methodological contribution is the formalisation of pedagogical strategy selection: multimodal evidence and spline-derived trajectory indicators inform an explicit, constrained algorithm for choosing between PS-I and I-PS under teacher oversight. Separating this decision from the subsequent selection of specialised agents makes the proposed decision process traceable, testable, and open to revision. The methodology thus provides a basis for examining when and why particular instructional strategies should be selected.

The retrospective findings motivate this approach but do not validate its effectiveness. Prospective studies must establish whether adaptive strategy selection improves independent mastery and delayed transfer relative to fixed instructional policies, while evaluating diagnostic reliability, teacher workload, and computational cost.

Declaration of Conflicting Interests

The author declares no potential conflicts of interest with respect to the research, authorship, and publication of this article.

Funding

The author received no specific financial support for the research, authorship, or publication of this article.

Ethical Considerations and Informed Consent

This retrospective study was conducted in accordance with the ethical recommendations of the International Innovation Centre. Participating Ukrainian educational institutions authorised the use of statistical data for research purposes. The educational activities underlying the study respected the principles of voluntary participation, informed consent, confidentiality, and responsible use of educational data.

Results are reported at the class-group level without disclosing individual participants' identities. The use and publication of handwritten student work, including mathematical solutions, diagrams, and written explanations, were restricted to examples covered by permission for research dissemination.

Data Protection and Privacy

The study followed principles of data minimisation, confidentiality, and protection of personal information.

Retrospective educational results are presented in aggregated form at the class-group level. Personally identifying information was excluded from the reported statistical analyses.

Where handwritten student-work examples are included, identifying details are removed or obscured to reduce the risk of identification. Permission to use educational records for research is distinguished from permission to reproduce individual handwritten materials in scientific publications.

The proposed EduHyperAgent framework incorporates privacy protection, teacher authorisation, controlled access to learning evidence, and human oversight as fundamental design requirements.

References

Asai, A., He, J., Shao, R., et al. (2026). Synthesizing scientific literature with retrieval-augmented language models. *Nature*. https://doi.org/10.1038/s41586-025-10072-4.

Baral, S., Botelho, A., Santhanam, A., Gurung, A., Cheng, L., & Heffernan, N. (2023). Auto-scoring student responses with images in mathematics. *Proceedings of the 16th International Conference on Educational Data Mining*, 362–369. https://doi.org/10.5281/zenodo.8115645.

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences, 122*(26), e2422633122. https://doi.org/10.1073/pnas.2422633122.

Bengio, Y., Clare, S., Prunkl, C., Murray, M., et al. (2026). *International AI safety report 2026* (DSIT 2026/001). https://internationalaisafetyreport.org/publication/international-ai-safety-report-2026.

Blikstein, P., & Worsley, M. (2016). Multimodal learning analytics and education data mining: Using computational technologies to measure complex learning tasks. *Journal of Learning Analytics, 3*(2), 220–238. https://doi.org/10.18608/jla.2016.32.11.

Bucalon, B., Yu, L., Shaw, T., Gašević, D., Kay, J., & Janssen, A. (2026). Applying generative AI to the learning analytics cycle to support competency development in health professions education. *Journal of Learning Analytics, 13*(2), 63–86. https://doi.org/10.18608/jla.2026.9145.

Bull, S., & Kay, J. (2016). SMILI: A framework for interfaces to learning data in open learner models, learning analytics and related fields. *International Journal of Artificial Intelligence in Education, 26*, 293–331. https://doi.org/10.1007/s40593-015-0090-8.

Campbell, M. K., Piaggio, G., Elbourne, D. R., & Altman, D. G. (2012). CONSORT 2010 statement: Extension to cluster randomised trials. *BMJ, 345*, e5661. https://doi.org/10.1136/bmj.e5661.

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823.

Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction, 4*, 253–278. https://doi.org/10.1007/BF01099821.

DigitalOcean. (2025, September 19). GRPO in reinforcement learning explained [Technical tutorial]. *LinkedIn Pulse*. https://www.linkedin.com/pulse/grpo-reinforcement-learning-explained-digitalocean-i4axc/.

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*(1), 64–71. https://doi.org/10.1007/s11528-014-0822-x.

Guo, D., Yang, D., Zhang, H., et al. (2025). DeepSeek-R1 incentivizes reasoning in LLMs through reinforcement learning. *Nature, 645*, 633–638. https://doi.org/10.1038/s41586-025-09422-z.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487.

He, Y., Yang, H., Lu, Y., Klimovic, A., & Alonso, G. (2025). Resource multiplexing in tuning and serving large language models. *Proceedings of the 2025 USENIX Annual Technical Conference*, 1639–1655. https://www.usenix.org/conference/atc25/presentation/he-yongjun.

Holmes, W., Porayska-Pomsta, K., Holstein, K., et al. (2022). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education, 32*, 504–526. https://doi.org/10.1007/s40593-021-00239-1.

Holstein, K., McLaren, B. M., & Aleven, V. (2019). Co-designing a real-time classroom orchestration tool to support teacher–AI complementarity. *Journal of Learning Analytics, 6*(2), 27–52. https://doi.org/10.18608/jla.2019.62.3.

International Innovative School. (n.d.). *Official website*. https://iis.org.ua/.

Kapur, M. (2014). Productive failure in learning Math. *Cognitive Science, 38*(5), 1008–1022. https://doi.org/10.1111/cogs.12107.

Khan Academy. (n.d.). Meet Khanmigo: Khan Academy's AI-powered teaching assistant and tutor. https://www.khanacademy.org/khan-labs.

Khosravi, H., Buckingham Shum, S., Chen, G., et al. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence, 3*, 100074. https://doi.org/10.1016/j.caeai.2022.100074.

Kortemeyer, G., Caspar, A., & Horica, D. (2025). Artificial-intelligence grading assistance for handwritten components of a calculus exam [Preprint]. *arXiv*. https://doi.org/10.48550/arXiv.2510.05162.

Kumar, A., Wongsirichot, T., & Nanthaamornphong, A. (2026). Generative AI in computing education: A systematic review and a framework for responsible integration. *Computers and Education: Artificial Intelligence*, 100677. https://doi.org/10.1016/j.caeai.2026.100677.

Létourneau, A., Deslandes Martineau, M., Charland, P., Karran, J. A., Boasen, J., & Léger, P. M. (2025). A systematic review of AI-driven intelligent tutoring systems (ITS) in K–12 education. *npj Science of Learning, 10*, Article 29. https://doi.org/10.1038/s41539-025-00320-7.

Li, X., & Abdul Razak, R. (2026). Toward an interpretive synthesis of personalization, co-design, and ethical mediation in

AI-supported higher education: An integrative review of agentic, generative, and adaptive tutoring systems. *Journal of Educational Computing Research* (online first). https://doi.org/10.1177/07356331261498975.

Lin, M. P.-C., Chang, D. H., Winne, P. H., Bowen, G. M., & Ryoo, J. (2026). Mapping generative AI in higher education teaching and learning: A scoping review. *International Journal of Educational Technology in Higher Education, 23*, Article 55. https://doi.org/10.1186/s41239-026-00630-9.

Liu, X., Tang, Q., Wu, W., Chen, X., & Wang, C. (2026). GenAI-agents in education: A systematic review of macro-level trends, educational roles, theoretical foundations and future agendas. *Interactive Learning Environments* (online first). https://doi.org/10.1080/10494820.2026.2672562.

Lu, C., Lu, C., Lange, R. T., Foerster, J., Clune, J., & Ha, D. (2024). The AI Scientist: Towards fully automated open-ended scientific discovery [Preprint]. *arXiv*. https://arxiv.org/abs/2408.06292.

Martin, R., & Johnson, S. (2023, July 12). Introducing NotebookLM. *The Keyword, Google*. https://blog.google/innovation-and-ai/technology/ai/notebooklm-google-ai/.

Microsoft. (n.d.). Content sources supported by Reading Coach. *Microsoft Support*. https://support.microsoft.com/en-us/education/learning-accelerators/content-sources-supported-by-reading-coach.

Molenaar, I. (2022). Towards hybrid human–AI learning technologies. *European Journal of Education, 57*(4), 632–645. https://doi.org/10.1111/ejed.12527.

Mu, S., Cui, M., & Huang, X. (2020). Multimodal data fusion in learning analytics: A systematic review. *Sensors, 20*(23), 6856. https://doi.org/10.3390/s20236856.

OpenAI. (2022, November 30). Introducing ChatGPT. https://openai.com/index/chatgpt/.

Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence, 2*, 100020. https://doi.org/10.1016/j.caeai.2021.100020.

Photomath. (n.d.). Photomath: Step-by-step math learning. https://photomath.com/.

Puech, R., Macina, J., Chatain, J., Sachan, M., & Kapur, M. (2025). Towards the pedagogical steering of large language models for tutoring: A case study with modeling productive failure. *Findings of ACL 2025*, 26291–26311. https://doi.org/10.18653/v1/2025.findings-acl.1348.

Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education, 26*, 582–599. https://doi.org/10.1007/s40593-016-0110-3.

Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.

Schick, T., Dwivedi-Yu, J., Dessì, R., et al. (2023). Toolformer: Language models can teach themselves to use tools. *Advances in Neural Information Processing Systems, 36*. https://papers.neurips.cc/paper_files/paper/2023/hash/d842425e4bf79ba039352da0f658a906-Abstract-Conference.html.

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms [Preprint]. *arXiv*. https://doi.org/10.48550/arXiv.1707.06347.

ScriptaGrad. (n.d.). *ScriptaGrad: AI-powered assessment of handwritten STEM work* [Software platform]. https://scripta.interschool.ch/.

Shao, Z., Wang, P., Zhu, Q., et al. (2024). DeepSeekMath: Pushing the limits of mathematical reasoning in open language models [Preprint]. *arXiv*. https://doi.org/10.48550/arXiv.2402.03300.

Sinha, T., & Kapur, M. (2021). When problem solving followed by instruction works: Evidence for productive failure. *Review of Educational Research, 91*(5), 761–798. https://doi.org/10.3102/00346543211019105.

Theis, M., Thomsen, M., Vogel, C., & Wasserfuhr, V. (2026). *The impact of intelligent tutoring systems in general education: A systematic review of causal evidence on learning outcomes, motivation, and socio-emotional effects*. European Commission. https://doi.org/10.2766/0766181.

Vachtsevanou, D., Ciortea, A., Mayer, S., & Lemée, J. (2023). Signifiers as a first-class abstraction in hypermedia multi-agent systems. *Proceedings of AAMAS 2023*, 1200–1208. https://www.ifaamas.org/Proceedings/aamas2023/pdfs/p1200.pdf.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems.

Educational Psychologist, 46(4), 197–221. https://doi.org/10.1080/00461520.2011.611369.

Wang, N., Zou, D., Xie, H., & Qin, S. J. (2026). A scoping review of generative AI-powered agentic AI in education: Research landscape, agentic capabilities, and insights from the frontier agent paradigm, exemplified by OpenClaw. *Computers and Education: Artificial Intelligence*, 100653. https://doi.org/10.1016/j.caeai.2026.100653.

Wolfram Research. (n.d.). *About Wolfram|Alpha: Making the world's knowledge computable*. https://www.wolframalpha.com/about.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x.

Woodward, J. (2026, July 16). NotebookLM is now Gemini Notebook. *The Keyword, Google*. https://blog.google/innovation-and-ai/products/gemini-notebook/notebooklm-gemini-notebook/.

Wooldridge, M. (2009). *An introduction to multiagent systems* (2nd ed.). Wiley.

Wu, Q., Bansal, G., Zhang, J., et al. (2024). AutoGen: Enabling next-gen LLM applications via multi-agent conversation. *Conference on Language Modeling (COLM)*. https://www.microsoft.com/en-us/research/publication/autogen-enabling-next-gen-llm-applications-via-multi-agent-conversation-framework/.

Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2023). ReAct: Synergizing reasoning and acting in language models. *International Conference on Learning Representations*. https://arxiv.org/abs/2210.03629.

Yao, X., Hu, Q., & Klimovic, A. (2025). DeltaZip: Efficient serving of multiple full-model-tuned LLMs. *Proceedings of EuroSys 2025*, 110–127. https://doi.org/10.1145/3689031.3717468.

Yu, Q., Zhang, Z., Zhu, R., et al. (2025). DAPO: An open-source LLM reinforcement learning system at scale. *Advances in Neural Information Processing Systems, 38*. https://doi.org/10.52202/085713-3775.

Zalizko, V. (2026a). Automated assessment of mathematical problem-solving in secondary school based on the integration of ChatGPT and MRIIA+. *Information Technologies and Learning Tools, 112*(2), 1–18. https://doi.org/10.33407/itlt.v112i2.6536.

Zalizko, V. (2026b). Spline-guided segmentation of handwritten physico-mathematical documents for improved OCR accuracy [Preprint]. *Research Square*. https://doi.org/10.21203/rs.3.rs-8471729/v1.