



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 4 No. 3 (2025) pp: 3877-3889

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Artificial Intelligence (AI) as the Reflective Partner: Empowering Teachers for Deep Learning Pedagogy

¹Abdul Rahman, ²Fitri Nurmahmudah, ³Eka Cahya Sari Putra, ⁴Arta Mulya Budi Harsono, ⁵Brezto Asagi Dewantara, ⁶M. Ziyah Takhqiqi Arsyad, ⁷Fitri Alfarisa

^{1,5}Teknologi Pendidikan, FKIP, Universitas Lambung Mangkurat

²Badan Riset dan Inovasi Nasional, Indonesia

^{3,4,6}PGSD, FKIP, Universitas Lambung Mangkurat

⁷PGSD, FKIP, Universitas Pendidikan Indonesia

abdulrahman@ulm.ac.id¹, fitri.nurmahmudah@brin.go.id², ekacsputra@ulm.ac.id³, artamulyabudi@ulm.ac.id⁴, brezto.tp@ulm.ac.id⁵, mzarsyad@ulm.ac.id⁶, alfarisa@upi.edu⁷

Abstract

As technology continues to advance, Artificial Intelligence (AI) is viewed as having great potential as a reflective partner for teachers in designing, implementing, and evaluating learning. However, the implementation of AI in the context of teacher reflection remains limited, partial, and not yet systematically integrated into professional development. This study aims to explore the role of AI as a reflective partner to empower teachers in implementing deep learning pedagogy. The method used is a systematic literature review (SLR) with a PRISMA approach. Data was obtained through article searches in the SpringerLink database using relevant keywords, then filtered based on methodological quality, topic suitability, and publication completeness. A total of 21 selected articles were qualitatively analyzed using Atlas.ti version 24 software to identify codes, categories, and research concept maps. The analysis results revealed five main themes of AI's role in deep learning pedagogy: (1) AI-Supported Deep Learning Lesson Design, which supports teachers in designing objectives, content, and anticipating learning barriers; (2) AI in Orchestrating Deep Learning Processes, which organizes classroom dynamics through inquiry prompts, multi-representation of concepts, and real-time formative feedback; (3) AI-Augmented Deep Learning Assessment, which strengthens authentic assessment through multi-level rubrics, evidence-based performance analysis, and verification of learning processes; (4) AI-Driven Reflective Teaching Practices, which facilitates data-driven reflection, action planning, and knowledge curation for professional development; and (5) Ethical and Contextual AI Integration, which emphasizes the importance of local context, ethical awareness, and data protection. Based on the findings, it is recommended that teachers receive practice-based training that not only focuses on technical skills but also on integrating AI into deep learning pedagogy principles. Governments and educational institutions should formulate policies that encourage the ethical, transparent, and contextually appropriate use of AI. Technology developers are advised to design AI as a co-designer and co-orchestrator while maintaining teachers as the main decision-makers. Further research is needed through longitudinal, cross-cultural, and cross-level studies, including in vocational education, to assess the long-term effectiveness of AI utilization in supporting teacher professionalism and achieving deep learning outcomes.

Keywords: Artificial Intelligence, Teacher Reflection, Deep Learning Pedagogy, Professional Development, Deep Learning.

1. Introduction

Improving teacher quality is a strategic mandate of Indonesia's national education system as stipulated in Law No. 14 of 2005 concerning Teachers and Lecturers, which requires the continuous development of pedagogical, personal, social, and professional competencies. According to the Regulation of the Director General of Teachers and Education Personnel No. 7607/B.B1/HK.03/2023, teachers are encouraged to improve learning practices that focus on students' cognitive activation through interactive activities. In addition, after conducting classroom observations, teachers are expected to reflect on their practices in response to notes and performance evaluation results as part of the cycle of instructional improvement. Regulation of the Ministry of Education, Culture, Research, and Technology No. 16 of 2022 on Process Standards emphasizes that learning processes must be designed to be interactive, inspiring, enjoyable, challenging, and motivating for students to actively participate. These characteristics are the essence of a student-centered deep learning approach.

Recent literature affirms that AI holds great potential as a reflective partner in developing teacher professionalism, particularly in observing, analyzing, and enhancing the quality of deep learning (Doğan et al.,

2025; Tammets & Ley, 2023). Previous research on pre-service teachers found that AI-powered dashboards can support teacher reflection with data-driven insights, transforming reflection from a traditional approach into a more systematic, evidence-based practice (Cai et al., 2025). Although AI's potential is widely acknowledged, its implementation to support reflection in deep learning remains limited. For example, the use of QuestionBot by pre-service teachers helps them recognize questioning patterns, but the trials remain small in scale (Sert, 2025). Generative AI enriches reflection through digital storytelling, yet its adoption is still dominant among pre-service teachers (Wei et al., 2025). Intelligent teaching analytics enhance collaborative reflection, but some teachers initially feel anxious about AI analysis (M. Wang et al., 2025). Teachers' trust in AI is influenced by AI literacy and cultural context, with significant skepticism still present (Viberg et al., 2024). Teacher questioning quality improved by 20% through AI feedback, but its effectiveness depends on teachers' trust in the accuracy of data (D. Demzky et al., 2025). Most AI-based professional development programs have not fully integrated deep learning pedagogy principles, tending to focus only on technical skills of AI use (Doğan et al., 2025). Teachers, therefore, need practice-based training to optimally utilize AI (Aljemely, 2024).

Normatively, teachers are expected to use AI to strengthen reflective practices within the deep learning approach, in accordance with national policies and educational research developments. However, in practice, AI implementation remains partial, limited to small trials, not integrated into ongoing professional development, and often not explicitly directed toward supporting deep learning pedagogy. This gap calls for research that connects scientific evidence and educational policies with concrete practices of using AI as a reflective partner for teachers. The urgency of this research lies in the need to bridge the gap between policy ideals, research findings, and practical realities. Amidst the accelerating integration of technology in education, Indonesia requires a framework for AI utilization aligned with teacher needs and oriented toward deep learning. This study is expected to contribute to the design of a contextual, sustainable model of teacher professional development that strategically integrates AI as a reflective partner. Based on the above background, this paper aims to explore AI as a reflective partner to empower teachers in implementing a deep learning approach.

2. Method

The design used in this article is a systematic literature review. The reason for using a systematic literature review is to obtain a comprehensive, objective, and accountable picture of AI as the Reflective Partner. The data collection technique employed was a literature search, selecting articles published in Scopus-indexed journals. The articles were retrieved through the SpringerLink database. The literature search method applied was PRISMA, which aims to obtain high-quality articles. This method helps with duplication removal, title and abstract screening, and assessing the relevance of articles to the research question or objectives. The procedure in this method consists of identification, screening, eligibility, and inclusion. The data analysis technique was qualitative, which involved the stages of data reduction, coding, and concept mapping. Data analysis was conducted with the assistance of Atlas.ti version 24 software. In detail, the procedure for determining the articles for analysis can be seen in Figure 1.

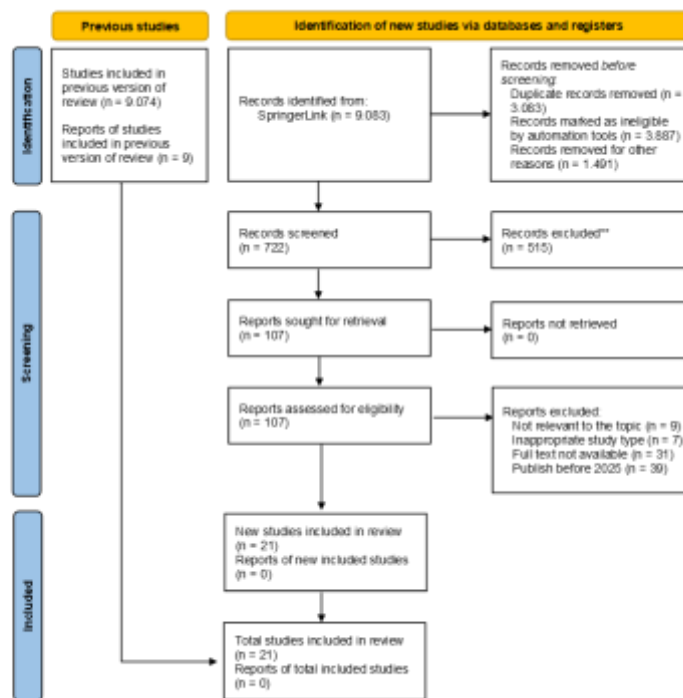


Figure 1. PRISMA AI as the Reflective Partner

Based on Figure 1, articles were identified and screened using keywords. The articles included in the analysis were identified based on the most recent publication years, specifically those published in globally indexed journals (Scopus) in 2025. Next, articles were searched using the following keywords: *deep learning pedagogy*, *AI integration in reflection*, *teacher reflection*, *learning reflection*, *learning pedagogy*, *deep learning approach*, *vocational school deep learning approach*, and *AI as the reflective partner*. The identified articles were then screened for relevance to this study's topic. Articles with strong methodology were prioritized. Additionally, the screening process considered the completeness of the article sections, such as IMRAD structure.

3. Result and Discussion

Result

Based on the identification and screening process conducted in the research method, 21 (twenty-one) articles were found to be highly relevant to this study. These articles were analyzed using the procedures presented in the methods section. The analysis of the 21 selected articles through a systematic literature review (SLR) revealed several major tendencies. First, in terms of publication period, most articles were published in 2025, indicating growing attention to the topic under investigation. This distribution shows the urgency and relevance of the issue within the current research context. Second, regarding research domains, the analyzed articles can be grouped into three main themes: (1) Conceptual and theoretical studies, which emphasize frameworks, models, or approaches; (2) Empirical studies, which highlight implementation, case studies, or practices in the field; and (3) Evaluative and comparative studies, which assess the effectiveness of certain approaches or evaluate their impact on learning outcomes/quality improvement. Third, in terms of research methodology, the majority of articles used qualitative approaches (such as case studies, interviews, or document analysis), followed by quantitative methods (surveys, experiments), and a smaller proportion adopted mixed methods. This pattern indicates that the issue under investigation is still widely explored through descriptive and exploratory approaches.

In addition, this review found that most articles emphasize implementation challenges such as resource limitations, teacher readiness, and policy support. Meanwhile, other articles highlight opportunities and recommendations such as the use of technology, capacity-building for educators, and collaboration with industry or related institutions. Overall, this SLR shows a general consensus that the researched topic has strategic relevance for both educational development and professional practice. However, there remains a research gap regarding long-term evaluation and quantitative impact measurement, which opens opportunities for further studies. The relevant articles analyzed in this study are presented in Table 1.

Table 1. Article Screening Result

No.	Author	Title
1	(Yu, 2024)	Research on the reform of university education pedagogy based on deep learning model in the context of information era
2	(Lin et al., 2025)	Enhancing reflective thinking in STEM education through experiential learning: The role of generative AI as a learning aid
3	(Engeness, 2025)	Cultural-historical perspective to design pedagogical ai for enhancing student writing
4	(Babaci-Wilhite & Motyka, 2025)	Teaching and learning for the common good: A case study on pedagogical engagement in the US higher education
5	(Rincón et al., 2025)	Instructional design for tutoring on interactive platforms: creating educational interventions overcoming the digital gap
6	(Jerasa & Ura, 2025)	Learning from tiktok: Quality and reach of #teachertok as a classroom management tool for teacher education
7	(Sager et al., 2025)	Farming with data: tracing critical tensions using data science for food justice
8	(van den Beuken et al., 2025)	Experiences of soft skills development and assessment by health sciences students and teachers: a qualitative study
9	(Prosen & Ličen, 2025)	Evaluating the digital transformation in health sciences education: a thematic analysis of higher education teachers' perspectives

10	(Tong et al., 2025)	The application of problem-based learning (PBL) guided by chatgpt in clinical education in the department of nephrology
11	(L. Wang et al., 2025)	Teaching innovation in a pharmacy course: integration of “questioning-training of comprehensive knowledge application” and a “teacher-ai-student interaction model”
12	(Gorham et al., 2025)	A microlearning app for peer feedback training and its effect on learning performance and self-confidence during an EFL speaking task
13	(X. Wang, 2024)	Deep learning-based personalized learning recommendation system design for "T++" guzheng pedagogy
14	(Breidenstein et al., 2025)	Exploring teaching quality from within: A praxeological perspective
15	(Z. Wang et al., 2025)	Feasibility study of using gpt for history-taking training in medical education: a randomized clinical trial
16	(Simon et al., 2025)	The postdigital classroom: a guide toward the imaginary
17	(Nolan & Zeidler, 2025)	Cultivating critical reflection through socioscientific perspectives
18	(Giskeødegård & Wahl, 2025)	Signature pedagogies of Norwegian simulator-based nautical training
19	(Asmawi & Alam, 2025)	A utopic techno-teacher coupling: designing an ad hoc activity map towards transformative pedagogical prototypes for reading instruction
20	(van den Berg, 2025)	Teachers’ experiences of using artificial intelligence from an open distance learning context: successes, challenges, and strategies for success
21	(Galeboe et al., 2025)	What is the impact of using design and technology pedagogy to support the attainment of 21st-century skills?

The analysis of the 21 articles, which served as the data sources for this research, was conducted using qualitative meaning analysis with the assistance of Atlas.ti version 24 software. Each article was deeply explored to uncover the meanings contained within, aligned with the research focus: *AI as the Reflective Partner: Empowering Teachers for Deep Learning Pedagogy*. The analysis process produced several codes and categories, which were then mapped into a concept map. The concept map illustrates the interrelationships between AI’s role and teachers’ reflective practices in supporting deep learning pedagogy. From this mapping, several key themes were identified (1) AI-Supported Deep Learning Lesson Design; (2) AI in Orchestrating Deep Learning Processes; (3) AI-Augmented Deep Learning Assessment; (4) AI-Driven Reflective Teaching Practices; and (5) Ethical and Contextual AI Integration.

1. AI-Supported Deep Learning Lesson Design

This theme highlights how AI supports teachers in designing deep learning more systematically, adaptively, and contextually. The analyzed articles show that AI helps teachers to:

- Identify students’ learning needs through learning data analysis.
- Provide recommendations for learning resources and teaching strategies that align with student profiles.
- Facilitate the design of project-based, collaborative, and problem-based learning that emphasizes higher-order thinking skills.

Thus, AI functions not only as a technical tool but as a co-designer, supporting teachers in creating richer deep learning experiences.

2. AI in Orchestrating Deep Learning Processes

This theme refers to AI’s role in coordinating and facilitating the flow of deep learning processes. Orchestration here includes managing activities, interactions, and classroom dynamics. AI enables:

- Real-time monitoring of student interactions to ensure active engagement.
- Adaptive interventions, such as suggesting additional activities or slowing the pace of instruction according to students’ needs.
- Support for teachers in managing complex learning, especially collaboration and reflective discussions.

Discussion

Based on the data analysis conducted with the assistance of Atlas.ti version 24, it can be explained that *AI as the Reflective Partner* consists of five main components. Each component contains distinct indicators, as follows:

AI-Supported Deep Learning Lesson Design

The findings from this component consist of AI-Enhanced Deep Learning Goals, Adaptive Content and Context Selection, and Anticipation of Learning Barriers.

AI-Enhanced Deep Learning Goals

AI can help teachers formulate measurable and evidence-based deep learning goals, for example, goals related to improving dialogic quality, student idea uptake, and higher-order questioning, by analyzing classroom discourse and providing specific feedback on teaching practices. Demszky's study shows that automated feedback increases the use of "focusing" questions that stimulate student reasoning (an indicator of higher-level cognitive engagement), with about a 20% improvement among teachers who received targeted recommendations (D. Demszky et al., 2025). Similar findings in earlier online contexts showed that AI can improve instructors' uptake of student ideas (Dorottya Demszky et al., 2023). The TalkMoves application maps teachers' talk moves (revoicing, probing, pressing for reasoning) and sparks reflection to enrich mathematical discussions, moving goals closer to "deep thinking" rather than rote memorization (Jacobs et al., 2022). In the area of goal design, the AI-dialogue analysis framework by Li et al. helps teachers link instructional goals with dialogue quality metrics (e.g., question quality, student idea engagement), thereby operationalizing deep learning goals (X. Li et al., 2025). In teacher education, Tammets & Ley (2023) emphasize the integration of AI in teacher noticing and professional vision, so that goal formulation is not merely a list of outcomes but the result of a "see-interpret-decide" process supported by data (Tammets & Ley, 2023). Moreover, generative AI can spark the formulation of reflective goals through digital storytelling that encourages causal analysis and alternative teaching strategies (Wei et al., 2025).

Based on the discussion above, it can be concluded that the establishment of deep learning goals can be guided by AI through: (a) discourse analytics to derive goal indicators (HOTS questions, student idea uptake), (b) teacher dashboards to monitor progress toward the goals, and (c) generative prompts to enrich the clarity and depth of reflective goals.

Adaptive Content and Context Selection

Once learning goals are clearly established, AI can support teachers in selecting and adaptively tailoring content and context based on classroom evidence (e.g., videos, transcripts, interaction patterns) and learners' needs. A design-based study by Wang et al. (2025) demonstrated that iteratively developed intelligent teaching analytics enhanced collaborative reflection and shared regulation; teachers used data to determine content foci, contextual examples, and instructional moments requiring reinforcement. In video-based learning environments, AI-powered dashboards can extract patterns from lesson recordings, thereby sharpening teachers' noticing and enabling more evidence-based decisions regarding content adaptation (Cai et al., 2025). The AI-dialogue framework proposed by Li et al. provides guidance for selecting activities that foster meaningful dialogue rather than merely adding content. Tammets & Ley (2023) highlight that when teachers are involved in the co-design of AI tools, content adaptations are more likely to align with curriculum standards, competency frameworks, and local contexts (Tammets & Ley, 2023). The finding Demszky dkk. (2023) are also pertinent: metrics such as teacher talk time and question variability can assist teachers in adapting activities to be more responsive to students' ideas (Dorottya Demszky et al., 2023).

Based on the foregoing analysis, it can be concluded that AI-supported content and context adaptation operates through: (a) learning analytics to map real-time needs, (b) video/discourse visualizations to guide activity adjustments, and (c) co-design processes to ensure that materials and activities remain relevant to local contexts and deep learning goals.

Anticipation of Learning Barriers

Well-designed lessons anticipate potential learning barriers. AI can predict and flag such barriers across three levels. (1) Practice level: indicators such as excessive teacher talk, limited wait-time, or shallow questioning can be automatically detected; targeted interventions (e.g., email feedback or micro-coaching) have been shown to improve questioning practices that stimulate reasoning (Demszky et al., 2025; Jacobs et al., 2022; Li et al., 2025); (2) Individual teacher level: teachers' trust in and literacy regarding AI strongly influence their willingness to adopt recommendations.; Viberg dkk. (2024) demonstrated that trust is shaped primarily by perceptions of usefulness, risk, and AI efficacy rather than demographic factors, making onboarding and ethical training crucial. (3) Professional development implementation level; Wang dkk. (2025) found that initial anxiety toward analytics can be mitigated through iterative design and collaborative support. Frameworks by Tammets & Ley (2023)

together with evidence on generative AI for reflection (Wei et al., 2025) emphasize the importance of scaffolding to move teacher reflection beyond descriptive habits. From these findings, it can be concluded that the anticipation of barriers requires: (a) early-warning metrics (e.g., questioning quality, idea uptake), (b) strengthening teacher trust and AI literacy through scaffolded ethical training, and (c) collaborative and iterative support to ensure that AI recommendations are genuinely integrated into classroom practice.

Synthesizing across relevant studies, it may be concluded that AI-Supported Deep Learning Lesson Design positions AI as a reflective partner in: (1) formulating operational and measurable deep learning goals through dialogue analytics and personalized feedback, (2) adaptively selecting and tailoring content and context using visualized classroom evidence, and (3) anticipating learning barriers at the practice, teacher, and PD implementation levels through early metrics, enhanced trust/literacy, and collaborative scaffolds. This approach integrates state-of-the-art evidence—ranging from improved questioning practices and student idea uptake to the maturation of professional reflection—thus advancing pedagogical designs that truly facilitate deep learning.

AI in Orchestrating Deep Learning Processes

The review identifies three key components of AI's orchestration role: AI-Generated Deep Inquiry Prompts, Multi-Representation Concept Delivery, and Real-Time Formative Feedback Support.

AI-Generated Deep Inquiry Prompts

Large Language Models (LLMs) and Generative AI can function as *Socratic partners*, stimulating higher-order, causal, and counterfactual questioning to drive student reasoning. A quasi-experimental study in engineering education showed that constructivist, inquiry-based prompting with ChatGPT enhanced conceptual understanding and reduced misconceptions compared to control groups (N = 120), underscoring AI's potential in structuring evidence-based inquiry (El Fathi et al., 2025). At the level of teaching praxis, Park and Choo (2024) in *JETS* outlined a structured model of prompt engineering for educators (persona, objective, source, constraints, evaluation), repositioning AI prompts from “answer-giving” toward “thinking-process activation.” In the arts, the LLaVA-Docent project demonstrated that multimodal LLMs trained via instruction-tuning can operate as personalized docents, generating tailored guiding questions for visual artifacts—an example of AI producing inquiry prompts across modalities to foster deeper appreciation and analysis (U. Lee et al., 2024).

Best practices for AI-Generated Inquiry Prompts include: (a) employing explicit pedagogical prompt frameworks (e.g., *JETS*) to align AI questioning with higher-order learning goals, (b) leveraging multimodal LLMs (e.g., LLaVA-Docent) to spark evidence-based dialogue around objects and representations, and (c) validating effectiveness through quasi-experimental designs in targeted subject contexts.

Multi-Representation Concept Delivery

Deep learning requires the orchestration of multiple representations (verbal, visual, symbolic, graphical, simulation-based). Evidence in chemistry education demonstrates that multimodal LLMs such as GPT-4V can “read” and reason with visual representations to support problem-solving, enabling co-explanation of text and image for abstract concepts (Alasadi & Baiz, 2024). Studies on LLaVA-Docent highlight cross-modal tutoring (text+image) in art appreciation, a transferable pattern for STEM subjects where teachers need multi-modal narration (description, sketching, highlighting) to link concepts, examples, and contexts (U. Lee et al., 2024). Theoretical position papers underscore the opportunities and challenges of LLM/MLLM orchestration across modalities (text, speech, image, video), stressing the need to integrate multi-representational strengths with learning design and cognitive load-sensitive evaluation (Küchemann et al., 2025). Moreover, experimental evidence on non-AI multi-representational tasks shows that performance improves when tasks demand inter-representational translation, providing a foundation for AI to optimize sequencing (e.g., diagram → equation → narrative explanation) (Hahn & Klein, 2023).

AI-enabled Multi-Representation Concept Delivery therefore operates through: (a) multimodal LLMs for co-explanation across text-image-simulation on difficult concepts, (b) sequencing designs guided by cognitive load evidence, and (c) assessment of understanding via inter-representational translation tasks.

Real-Time Formative Feedback Support

Deep learning is contingent upon real-time formative feedback that is specific, actionable, and supportive of self-regulation. Randomized controlled trials in K–12 settings revealed that automated micro-coaching (via email/dashboard) increased teachers' use of focusing questions by approximately 20% over controls, illustrating AI's role in modulating teaching practices mid-instruction (D. Demszky et al., 2025). At scale, a large RCT in online courses (n ≈ 1,136 instructors) demonstrated that AI feedback improved instructor uptake of student ideas (~+13%), confirming AI's contribution to enriching teacher-student dialogue (Dorotyya Demszky et al., 2023). In mathematics classrooms, TalkMoves provides automated feedback on discourse patterns (revoicing, probing, pressing for reasoning), enhancing reflective practice and enriching discussion quality (Jacobs et al., 2022;

Scornavacco, 2022). The CEAI perspective suggests that LLM can bring evidence-based feedback into the classroom in a more timely manner, positively correlating with student cognitive and motivational outcomes—while emphasizing the need for human supervision and task context (Meyer et al., 2024). argue that LLMs can deliver more timely evidence-based feedback correlated with cognitive and motivational outcomes, while emphasizing human oversight. Recent work also highlights modality of feedback delivery (AI-peer-instructor blends), accuracy, and trust as critical adoption factors (Burner et al., 2025; Venter et al., 2025).

Effective Real-Time Formative Feedback Support includes: (a) alignment of dialogic metrics with higher-order thinking goals (e.g., focusing questions, idea uptake), (b) timely, actionable feedback through dashboards/emails, and (c) human-in-the-loop safeguards to ensure accuracy, fairness, and trust. Overall, AI in Orchestrating Deep Learning Processes positions AI as a *co-orchestrator* that: (1) generates inquiry prompts aligned with higher-order thinking, (2) adaptively delivers multi-representational content sequences to facilitate inter-modal translation, and (3) provides real-time formative feedback shown in RCTs and quasi-experiments to improve classroom dialogue and practice.

AI-Augmented Deep Learning Assessment

This component comprises three pillars: AI-Crafted Multi-Level Rubrics, Evidence-Based Performance Analysis, and Verification of the Original Learning Process.

AI-Crafted Multi-Level Rubrics

LLMs are increasingly capable of designing or enriching multi-level rubrics and anchor examples aligned with higher-order outcomes. Experimental studies show that explicit criterion-based modeling improves rubric alignment and scoring reliability, even for free-access models, provided the rubric criteria are detailed and contextualized (e.g., HOTS indicators, argument quality, evidence use) (Zhang et al., 2024). Automated scoring improves when Chain-of-Thought reasoning is applied, yielding more consistent open-response ranking and transparent criterion-based justifications (G.-G. Lee et al., 2024). Collaborative grading approaches (LLM + human) enhance fairness and explainability (Win Myint et al., 2024). Systematic reviews further highlight efficiency gains while maintaining the necessity of human oversight (J. Li et al., 2024). A recent study in Assessment & Evaluation in Higher Education evaluated a GenAI-based rubric model embedded with motivational design features and behavioral framing; it concluded that it was “reliable but supervised,” suitable for use within a clear human control framework (Askari, 2025). These findings suggest that AI-generated, multilevel rubrics perform optimally when domain-rich criteria are available, anchor examples are available, and double-marking procedures are in place for calibration. Effective AI-crafted rubrics require: (a) domain-rich, detailed criteria, (b) explainable scoring through CoT or criterion-based justifications, and (c) human-in-the-loop double-marking and calibration processes.

Evidence-Based Performance Analysis,

Deep learning assessment depends on analyzing performance evidence—artifacts, process traces, argumentative reasoning, and actionable feedback. GPT-4 has shown effectiveness in diagnosing writing strengths/weaknesses and producing feedback on structure, cohesion, and use of evidence (Dai et al., 2024; Pack et al., 2024). LLMs combined with comparative judgment approaches improve essay scoring reliability, minimizing rubric bias (HCI/ACM, 2024). Classroom-level trials confirm that AI-generated feedback improves questioning practices and idea uptake, strengthening evidence quality for reasoning and making performance analysis more robust (G.-G. Lee et al., 2024). At the classroom level, automated feedback to teachers (question-answer quality metrics, concrete examples) increased the use of focusing questions that encourage student reasoning (~+20% in K-12 RCTs), resulting in improved quality of performance evidence (dialogue/reasoning) and easier analysis (Demszky et al., 2025; Large-scale RCTs on online courses also increase student idea uptake by ~+13%) (Dorottya Demzky et al., 2023). Overall, the findings demonstrate that artifact-based and discourse-based analytics—combined with rubric metrics and feedback loops—can measure progress toward deep learning outcomes (reasoning, evidence use, and transfer). Evidence-Based Performance Analysis combines: (a) criterion-based scoring (rubric/CoT/CJ); (b) action-specific feedback; and (c) reasoning evidence metrics (e.g., focusing questions, idea uptake) to assess depth of learning—not just final accuracy.

Verification of Original Learning Process

In the GenAI era, verifying the originality of learning processes is central to assessment validity. AI detectors are currently unreliable and should not serve as sole evidence of integrity (Elkhatat et al., 2023). stronger approaches include process-based assessment through keystroke logging, revision-flow mining, and interaction traces. Empirical studies confirm that keystroke dynamics can distinguish bona fide from AI-assisted writing (LREC 2024; Electronics 2024; Pattern Recognition 2025 challenge). In addition, studies of sensemaking/writing processes and process tracing confirm that recording the process (reading-writing-revising) provides a strong

indicator of originality and depth of thinking (Tarchi et al., 2024; Thomas et al., 2024). In practice, verification is strengthened with a process portfolio + AI usage statement + short viva/directed observation; if necessary, multi-detector aggregation can reduce false positives, but must still be combined with process evidence and human assessment (Hyatt et al., 2025). Robust Verification of Original Learning Process includes: (a) evidence of process (keystrokes/revision logs/citations); (b) transparency of AI use + viva/observation; (c) use of AI detectors only as supporting indicators, not the sole determinant.

Based on relevant studies, it can be concluded that AI-Augmented Deep Learning Assessment positions AI as a reinforcement of assessment validity and depth through three pillars: (1) designing a criteria-rich and explainable multi-level rubric with human oversight; (2) evidence-based performance analysis that combines criteria-based scoring, actionable feedback, and reasoning/argumentation metrics; and (3) verification of the originality of the learning process based on process traces, portfolios, and viva; AI detectors are only complementary. Designed in this way, the assessment ecosystem not only assesses results, but also the deep thinking processes that are at the heart of deep learning.

AI-Driven Reflective Teaching Practices

This component encompasses AI-Supported Action Planning, Data-Driven Reflection Summaries, and Knowledge Curation for Professional Growth.

AI-Supported Action Planning

AI is increasingly capable of supporting teachers in developing post-reflection action plans that are targeted toward instructional improvement. The study *Integrating AI Tools in Teacher Professional Learning* by Tammets and Ley (2023) proposed a model in which AI assists teachers in identifying improvement priorities, mapping concrete steps, and linking them to relevant learning resources. Demszky et al. (2025), teachers who received automated feedback from classroom recordings were able to use it to set instructional change targets, such as increasing the proportion of focusing questions to stimulate student reasoning. Sert (2025) demonstrated that automatic question detection enabled pre-service teachers to formulate more specific action plans for improving the quality of classroom questioning. Wang et al. (2025) further showed that intelligent teaching analytics can guide teaching teams in collaboratively designing instructional strategies based on classroom data. Similarly, Wei et al. (2025) highlighted how generative AI can suggest improvement steps after teachers engage in reflective digital storytelling. AI-supported action planning is effective when: (a) AI-generated feedback is translated into specific instructional targets, (b) AI recommendations are adapted to the local classroom and teacher context, and (c) follow-up cycles are established to monitor the realization of action plans.

Data-Driven Reflection Summaries

Data-driven reflection summaries enable teachers to gain a holistic understanding of their instructional practices. Cai et al. (2024) showed that AI-powered teacher dashboards analyzing classroom videos can produce summaries of interaction patterns, guiding teachers in identifying strengths and areas for improvement. Jacobs et al. (2022) through the *TalkMoves* tool, found that visualizing discourse moves facilitated the development of richer reflective narratives. Demszky et al. (2023) reported that presenting metrics such as teacher talk time and student idea uptake in concise reports enhanced teachers' awareness of classroom dynamics. Viberg et al. (2024) added a critical dimension: teachers are more likely to rely on AI-generated summaries when they trust the quality and relevance of the insights. Sert (2025) further noted that pre-service teachers who were initially hesitant ultimately came to depend on AI summaries to guide reflective discussions with mentors. Data-driven reflection summaries are most effective when: (a) they include pedagogically relevant key metrics, (b) they are presented through clear visualizations, and (c) they are accompanied by contextual explanations to avoid misinterpretation.

Knowledge Curation for Professional Growth

AI can curate knowledge resources—including scholarly articles, exemplary practices, and model teaching videos—aligned with teachers' professional development needs emerging from reflection. Tammets and Ley (2023) emphasized that involving teachers in the co-design of curation systems ensures the recommendations are more usable. Wang et al. (2025) showed that linking teaching analytics to repositories of proven practices accelerates teachers' access to effective instructional solutions. In their systematic review, Doğan et al. (2025) found that effective AI-focused professional development programs often include resource recommendation modules tailored to teacher profiles. Wei et al. (2025) illustrated how generative AI can suggest theoretical references or alternative strategies during creative reflective processes. Sert (2025) underscored AI's potential role as a *knowledge broker*, connecting teacher reflections with empirical evidence or strategies from the latest educational research.

Based on relevant studies, it can be concluded that AI-Driven Reflective Teaching Practices positions AI as a catalyst in the teacher reflection cycle through: (1) support for measurable and context-relevant action planning, (2) provision of data-driven reflection summaries that enrich pedagogical awareness, and (3) knowledge

uration that connects reflection with quality learning resources for continuous professional development. The combination of these three components forms a more structured, evidence-based, and adaptive reflection ecosystem, thereby accelerating the improvement of teaching quality.

Ethical and Contextual AI Integration

The review highlights three core dimensions of ethical and contextual AI integration: local contextualization and technical readiness, ethical awareness and bias sensitivity, and data privacy under the minimal-data principle.

Local Contextualization and Technical Readiness

Effective AI integration in education requires attention to local contexts—language, culture, and policy—as well as institutional technical readiness. Viberg et al. (2024) found that teachers' trust in AI is strongly shaped by cultural and workplace factors; teachers in countries with mature educational AI policies were more prepared to adopt such systems. Tammets and Ley (2023) emphasized the importance of co-design between developers and local educators to ensure alignment with national curricula and cultural values. Doğan et al. (2025) highlighted that technical readiness—including network infrastructure, device access, and teachers' digital competence—is a prerequisite for successful AI use in professional development. Cai et al. (2024) showed that adapting AI interfaces and instructional language to the local context improved pre-service teachers' engagement in video-based learning. Sert (2025) further observed that AI systems adjusting terminology and analytic styles to local practices gained greater teacher acceptance. Local contextualization and readiness require: (a) adaptation to local language, curriculum, and values, (b) strengthening of infrastructure and digital literacy, and (c) co-design processes to ensure relevance and acceptance.

Ethical Awareness and Bias Sensitivity

AI in education must be integrated with ethical awareness and sensitivity to bias to avoid reinforcing inequities. Aljemely (2024) found that teachers' reluctance to adopt AI often stemmed from concerns about bias, accuracy, and ethical implications. Viberg et al. (2024) stressed that ethical AI training, which builds teachers' understanding of risks and student impacts, can increase trust. Wei et al. (2025) cautioned that generative AI for reflection must be supervised to avoid introducing stereotypes or inaccurate assumptions about teaching practices. L'Enfant (2024) demonstrated that AI as a reflective coach in language education heightened teachers' awareness of data ethics and the need to critically verify AI recommendations. Doğan et al. (2025) recommended embedding bias detection modules and ethical AI literacy within teacher professional development programs. Ethical awareness and bias sensitivity require: (a) structured ethical and bias training, (b) human verification of AI-generated recommendations, and (c) institutional policies ensuring fairness and equity.

Data Privacy and Minimal Data Principle

Data privacy is a fundamental principle of responsible AI integration. According to the Pitt Teaching Center (2025), educational AI must adhere to the principle of minimal data collection, gathering only what is strictly necessary. Wang et al. (2025), in their study on intelligent teaching analytics, stressed the importance of transparent data governance, including clear communication with teachers regarding the types and purposes of collected data. Tammets and Ley (2023) recommended AI systems that store data locally or anonymously to reduce privacy risks. Viberg et al. (2024) noted that unclear privacy policies significantly undermine teacher trust. Elkhayat et al. (2023) warned that plagiarism detection and AI-generated content identification must comply with privacy regulations and avoid accessing student data without explicit consent. Data privacy and minimal data practices entail: (a) policies limiting data collection to essentials, (b) transparency in data-use purposes, and (c) safeguards such as anonymization, encryption, and local storage wherever possible.

Based on the review, Ethical and Contextual AI Integration requires three foundational pillars: (1) local contextualization and technical readiness to ensure alignment with culture, curriculum, and infrastructure, (2) ethical awareness and bias sensitivity to safeguard fairness and accuracy, and (3) robust data protection under the minimal-data principle. Together, these dimensions form the foundation for AI integration that is not only pedagogically effective but also socially and ethically responsible.

The novelty of this research lies in positioning AI as a *reflective partner* rather than a mere technical tool. This reframing introduces a new perspective in the literature on deep learning: teachers are not only leveraging AI for efficiency but also for empowering reflective practice and enriching pedagogy. Thus, this study opens new scholarly discourse on how AI integration can strengthen teacher professionalism while enhancing the quality of instructional practice in 21st-century education.

4. Conclusion

The findings of this study indicate that Artificial Intelligence (AI) has significant potential as a reflective partner for teachers in implementing deep learning pedagogy. The analysis of 21 recent articles highlights five major contributions of AI: (1) supporting adaptive deep learning lesson design, (2) orchestrating learning processes in real time, (3) strengthening evidence-based and authentic assessment, (4) enriching teacher reflection practices through data and knowledge curation, and (5) ensuring ethical and contextual AI integration. The novelty of this research lies in the understanding of AI not merely as a technical tool, but as a reflective partner that strengthens teachers' capacity for continuous improvement and enhances the quality of deep learning. Based on the research findings, teachers need practice-based training that goes beyond technical skills and emphasizes AI integration into the principles of deep learning pedagogy. Governments and educational institutions are encouraged to establish policies that support transparent, ethical AI use aligned with national curricula and local contexts, including infrastructure, culture, and teachers' digital literacy. Technology developers should design AI as a co-designer and co-orchestrator of learning, while keeping teachers as the primary decision-makers through a human-in-the-loop approach. Teachers and schools are advised to gradually integrate AI platforms into data-driven reflection and formative assessment, and to build collaborations with research institutions and technology industries so that AI adoption becomes more relevant and sustainable. This article also opens avenues for practical implementation and further research. At the implementation level, AI can be applied in teacher dashboards, adaptive assessment, and data-based reflection systems in schools. At the research level, longitudinal studies are needed to evaluate the long-term impact of AI on teacher professionalism and students' deep learning achievements, as well as quantitative experiments to test AI effectiveness in real classrooms. Furthermore, cross-cultural and cross-level research—including vocational schools—is crucial to understand the social, ethical, and technical factors influencing AI adoption. Thus, this research provides a foundational framework for the development of an AI-based teacher professional development model that is contextual, sustainable, and ethical, while bridging the gap between policy, research, and educational practice.

Referensi

- Alasadi, E. A., & Baiz, C. R. (2024). Multimodal Generative Artificial Intelligence Tackles Visual Problems in Chemistry. *Journal of Chemical Education*, 101(7), 2716–2729. <https://doi.org/10.1021/acs.jchemed.4c00138>
- Aljemely, Y. (2024). Challenges and best practices in training teachers to utilize artificial intelligence: a systematic review. *Frontiers in Education*, 9(October), 1–10. <https://doi.org/10.3389/educ.2024.1470853>
- Asmawi, A., & Alam, M. S. (2025). A utopic techno-teacher coupling: designing an ad hoc activity map towards transformative pedagogical prototypes for reading instruction. *Discover Education*, 4(1). <https://doi.org/10.1007/s44217-025-00585-5>
- Babaci-Wilhite, Z., & Motyka, M. (2025). Teaching and learning for the common good: A case study on pedagogical engagement in the US higher education. *Prospects*. <https://doi.org/10.1007/s11125-025-09733-z>
- Breidenstein, G., Heinzl, F., Schreyer, P., & Tyagunova, T. (2025). Exploring teaching quality from within: A praxeological perspective. *Zeitschrift Für Grundschulforschung*, 18(2), 469–482. <https://doi.org/10.1007/s42278-025-00239-3>
- Burner, T., Lindvig, Y., & Wærmess, J. I. (2025). “We Should Not Be Like a Dinosaur”—Using AI Technologies to Provide Formative Feedback to Students. *Education Sciences*, 15(1), 1–13. <https://doi.org/10.3390/educsci15010058>
- Cai, H., Lu, L., Han, B., Wong, L. H., & Gu, X. (2025). Exploring pre-service teachers' reflection mediated by an AI-powered teacher dashboard in video-based professional learning: a pilot study. *Educational Technology Research and Development*, 73(2), 1129–1154. <https://doi.org/10.1007/s11423-024-10442-1>
- Dai, W., Tsai, Y.-S., Lin, J., Aldino, A., Jin, H., Li, T., Gašević, D., & Chen, G. (2024). Assessing the proficiency of large language models in automatic feedback generation: An evaluation study. *Computers and Education: Artificial Intelligence*, 7, 100299. <https://doi.org/https://doi.org/10.1016/j.caeai.2024.100299>
- Demszky, D., Liu, J., Hill, H. C., Sanghi, S., & Chung, A. (2025). Automated feedback improves teachers' questioning quality in brick-and-mortar classrooms: Opportunities for further enhancement. *Computers & Education*, 227(1), 105183. <https://doi.org/https://doi.org/10.1016/j.compedu.2024.105183>
- Demszky, Dorotya, Liu, Jing, Hill, Heather C, Jurafsky, Dan, & Piech, Chris. (2023). Can Automated Feedback Improve Teachers' Uptake of Student Ideas? Evidence From a Randomized Controlled Trial in a Large-Scale Online Course. *Educational Evaluation and Policy Analysis*, 46(3), 483–505. <https://doi.org/10.3102/01623737231169270>
- Doğan, S., Y. Nalbantoglu, Ü., Celik, I., & Dogan, N. (2025). Artificial intelligence professional development: a systematic review of TPACK, designs, and effects for teacher learning. *Professional Development in Education*, 51(1), 1–28. <https://doi.org/10.1080/19415257.2025.2454457>
- El Fathi, T., Saad, A., Larhizil, H., Lamri, D., & Al Ibrahim, E. M. (2025). Integrating generative AI into STEM education: enhancing conceptual understanding, addressing misconceptions, and assessing student acceptance. *Disciplinary and Interdisciplinary Science Education Research*, 7(1), 1–21. <https://doi.org/10.1186/s43031-025-00125-z>
- Elkhatat, A. M., Elsaid, K., & Almeer, S. (2023). Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. *International Journal for Educational Integrity*, 19(1), 17. <https://doi.org/10.1007/s40979-023-00140-5>
- Engeness, I. (2025). Cultural-Historical Perspective to Design Pedagogical AI for Enhancing Student Writing. *Technology, Knowledge and Learning*, 2(1). <https://doi.org/10.1007/s10758-025-09876-0>
- Galeboe, K. A., Moalosi, R., Rapitsenyane, Y., & Ruele, V. (2025). What is the impact of using design and technology pedagogy to support the attainment of 21st-century skills? *Discover Education*, 4(1). <https://doi.org/10.1007/s44217-025-00604-5>
- Giskeødegård, M. F., & Wahl, A. M. (2025). Signature pedagogies of Norwegian simulator-based nautical training. *Discover Education*, 4(1). <https://doi.org/10.1007/s44217-025-00523-5>
- Gorham, T., Majumdar, R., & Ogata, H. (2025). A microlearning app for peer feedback training and its effect on learning performance and self-confidence during an EFL speaking task. *Smart Learning Environments*, 12(1). <https://doi.org/10.1186/s40561-025-00387-0>

- Hahn, L., & Klein, P. (2023). The impact of multiple representations on students' understanding of vector field concepts: Implementation of simulations and sketching activities into lecture-based recitations in undergraduate physics. *Frontiers in Psychology, 13*. <https://doi.org/10.3389/fpsyg.2022.1012787>
- Hyatt, J. P. K., Bienenstock, E. J., Firetto, C. M., Woods, E. R., & Comus, R. C. (2025). Using aggregated AI detector outcomes to eliminate false positives in STEM student writing. *Advances in Physiology Education, 49*(2), 486–495. <https://doi.org/10.1152/advan.00235.2024>
- Jacobs, J., Scornavacco, K., Hartly, C., Suresh, A., Lai, V., & Sumner, T. (2022). Promoting rich discussions in mathematics classrooms: Using personalized, automated feedback to support reflection and instructional change. *Teaching and Teacher Education, 112*(1), 1–13. <https://doi.org/10.1016/j.tate.2022.103631>
- Jerasa, S., & Ura, S. K. (2025). Learning from tiktok: Quality and reach of #teachertok as a classroom management tool for teacher education. *TechTrends*. <https://doi.org/10.1007/s11528-025-01098-6>
- Küchemann, S., Avila, K. E., Dinc, Y., Hortmann, C., Revenga, N., Ruf, V., Stausberg, N., Steinert, S., Fischer, F., Fischer, M., Kasneci, E., Kasneci, G., Kuhr, T., Kutyniok, G., Malone, S., Sailer, M., Schmidt, A., Stadler, M., Weller, J., & Kuhn, J. (2025). On opportunities and challenges of large multimodal foundation models in education. *Npj Science of Learning, 10*(1), 11. <https://doi.org/10.1038/s41539-025-00301-w>
- Lee, G.-G., Latif, E., Wu, X., Liu, N., & Zhai, X. (2024). Applying large language models and chain-of-thought for automatic scoring. *Computers and Education: Artificial Intelligence, 6*, 100213. <https://doi.org/https://doi.org/10.1016/j.caeai.2024.100213>
- Lee, U., Jeon, M., Lee, Y., Byun, G., Son, Y., Shin, J., Ko, H., & Kim, H. (2024). LLaVA-docent: Instruction tuning with multimodal large language model to support art appreciation education. *Computers and Education: Artificial Intelligence, 7*, 100297. <https://doi.org/https://doi.org/10.1016/j.caeai.2024.100297>
- L'Enfant, J. (2024). AI as a reflective coach in graduate ESL practicum: activity theory insights into student-teacher development. *European Journal of Open, Distance and E-Learning, 26*(1), 1–19. <https://doi.org/10.2478/eurodl-2024-0003>
- Li, J., Jangamreddy, N., Bhansali, R., Hisamoto, R., Zaphir, L., Dyda, A., & Glencross, M. (2024). AI-assisted marking: Functionality and limitations of ChatGPT in written assessment evaluation. *Australasian Journal of Educational Technology, 40*(4), 56–72. <https://doi.org/10.14742/ajet.9463>
- Li, X., Han, G., Fang, B., & He, J. (2025). Advancing the In-Class Dialogic Quality: Developing an Artificial Intelligence-Supported Framework for Classroom Dialogue Analysis. *Asia-Pacific Education Researcher, 34*(1), 495–509. <https://doi.org/10.1007/s40299-024-00872-z>
- Lin, C. J., Lee, H. Y., Wang, W. S., Huang, Y. M., & Wu, T. T. (2025). Enhancing reflective thinking in STEM education through experiential learning: The role of generative AI as a learning aid. *Education and Information Technologies, 30*(5), 6315–6337. <https://doi.org/10.1007/s10639-024-13072-5>
- Meyer, J., Jansen, T., Schiller, R., Liebenow, L. W., Steinbach, M., Horbach, A., & Fleckenstein, J. (2024). Using LLMs to bring evidence-based feedback into the classroom: AI-generated feedback increases secondary students' text revision, motivation, and positive emotions. *Computers and Education: Artificial Intelligence, 6*, 100199. <https://doi.org/https://doi.org/10.1016/j.caeai.2023.100199>
- Nolan, E., & Zeidler, D. (2025). Cultivating critical reflection through socioscientific perspectives. *Disciplinary and Interdisciplinary Science Education Research, 7*(1). <https://doi.org/10.1186/s43031-025-00136-w>
- Pack, A., Barrett, A., & Escalante, J. (2024). Large language models and automated essay scoring of English language learner writing: Insights into validity and reliability. *Computers and Education: Artificial Intelligence, 6*, 100234. <https://doi.org/https://doi.org/10.1016/j.caeai.2024.100234>
- Park, Jiyeon, & Choo, Sam. (2024). Generative AI Prompt Engineering for Educators: Practical Strategies. *Journal of Special Education Technology, 40*(3), 411–417. <https://doi.org/10.1177/01626434241298954>
- Pittsburgh, U. (2025, January 24). *Generative AI: Encouraging Academic Integrity*. https://teaching.pitt.edu/resources/encouraging-academic-integrity/?utm_source=chatgpt.com
- Prosen, M., & Ličen, S. (2025). Evaluating the digital transformation in health sciences education: a thematic analysis of higher education teachers' perspectives. *BMC Medical Education, 25*(1). <https://doi.org/10.1186/s12909-025-07420-3>
- Rincón, Y. R., Munárriz, A., Campión Arrastia, M. J., & Goicoechea López-Vailo, M. I. (2025). Instructional design for tutoring on interactive platforms: creating educational interventions overcoming the digital gap. *Educational Technology Research and Development*. <https://doi.org/10.1007/s11423-025-10516-8>
- Sager, M. T., Sherard, M. K., & Petrosino, A. J. (2025). Farming with data: tracing critical tensions using data science for food justice. *Journal of Computing in Higher Education*. <https://doi.org/10.1007/s12528-025-09466-0>
- Scornavacco, K. (2022). *Automated Feedback on Discourse Moves: Teachers' Perceived Utility of a Big Data Tool*. April, 1–36. <https://doi.org/10.3102/1887987>
- Sert, O. (2025). Partnering with AI in teacher education? Using an automatic question detection tool to reflect on classroom interaction. *Journal of Research on Technology in Education, 0*(0), 1–19. <https://doi.org/10.1080/15391523.2025.2504355>
- Simon, S., Vassaux, C., Guigon, G., Čarapina, M., & Consoli, T. (2025). The postdigital classroom: a guide toward the imaginary. *Postdigital Science and Education*. <https://doi.org/10.1007/s42438-025-00565-w>
- Tammets, K., & Ley, T. (2023). Integrating AI tools in teacher professional learning: a conceptual model and illustrative case. *Frontiers in Artificial Intelligence, 6*(1), 1–10. <https://doi.org/10.3389/frai.2023.1255089>
- Tarchi, C., Villalón, R., Vandermeulen, N., Casado-Ledesma, L., & Fallaci, A. P. (2024). Recursivity in source-based writing: a process analysis. *Reading and Writing, 37*(10), 2571–2593. <https://doi.org/10.1007/s11145-023-10482-8>
- Thomas, N., Hultgren, A. K., Zuaro, B., Yuksel, D., Wingrove, P., Nao, M., & Beach, D. (2024). Process Tracing for applied linguistics. *Research Methods in Applied Linguistics, 3*(2), 100118. <https://doi.org/https://doi.org/10.1016/j.rmal.2024.100118>
- Tong, X., Hu, Y., Long, Y., Zhang, Q., Yang, Y., Yuan, J., & Zha, Y. (2025). The application of problem-based learning (PBL) guided by chatgpt in clinical education in the department of nephrology. *BMC Medical Education, 25*(1). <https://doi.org/10.1186/s12909-025-07427-w>
- van den Berg, G. (2025). Teachers' experiences of using artificial intelligence from an open distance learning context: successes, challenges, and strategies for success. *Discover Education, 4*(1). <https://doi.org/10.1007/s44217-025-00596-2>
- van den Beuken, M., Loos, I., Maas, E., Stunt, J., & Kuijper, L. (2025). Experiences of soft skills development and assessment by health sciences students and teachers: a qualitative study. *BMC Medical Education, 25*(1). <https://doi.org/10.1186/s12909-025-07289-2>
- Venter, J., Coetzee, S. A., & Schmulian, A. (2025). Exploring the use of artificial intelligence (AI) in the delivery of effective feedback. *Assessment and Evaluation in Higher Education, 50*(4), 516–536. <https://doi.org/10.1080/02602938.2024.2415649>
- Viberg, O., Cukurova, M., Feldman-Maggor, Y., Alexandron, G., Shirai, S., Kanemune, S., Wasson, B., Tømte, C., Spikol, D., Milrad, M., Coelho, R., & Kizilcec, R. F. (2024). What Explains Teachers' Trust in AI in Education Across Six Countries? *International Journal of Artificial Intelligence in Education, 1*(1), 1–29. <https://doi.org/10.1007/s40593-024-00433-x>

- Wang, L., Xu, Y., Zhang, M., Bai, R., & Xie, T. (2025). Teaching innovation in a pharmacy course: integration of “questioning-training of comprehensive knowledge application” and a “teacher-ai-student interaction model.” *BMC Medical Education*, 25(1). <https://doi.org/10.1186/s12909-025-07549-1>
- Wang, M., Chen, Z., Xu, Y., Maheshi, B., & Gašević, D. (2025). Intelligent teaching analytics for collaborative reflection: investigating pre-service teachers’ perceptions, experiences and shared regulation processes. *International Journal of Educational Technology in Higher Education*, 22(1), 1–25. <https://doi.org/10.1186/s41239-025-00538-w>
- Wang, X. (2024). Deep learning-based personalized learning recommendation system design for “T++” guzheng pedagogy. *International Journal of Information Technology (Singapore)*, 16(5), 2775–2781. <https://doi.org/10.1007/s41870-024-01871-5>
- Wang, Z., Fan, T. T., Li, M. L., Zhu, N. J., & Wang, X. C. (2025). Feasibility study of using gpt for history-taking training in medical education: a randomized clinical trial. *BMC Medical Education*, 25(1). <https://doi.org/10.1186/s12909-025-07614-9>
- Wei, X., Wang, L., Koszalka, T. A., Lee, L.-K., & Liu, R. (2025). Enhancing pre-service teachers’ reflective thinking skills through generative AI-assisted digital storytelling creation: A three-dimensional framework analysis. *Computers & Education*, 235(1), 105356. <https://doi.org/https://doi.org/10.1016/j.compedu.2025.105356>
- Win Myint, P. Y., Lo, S. L., & Zhang, Y. (2024). Harnessing the power of AI-instructor collaborative grading approach: Topic-based effective grading for semi open-ended multipart questions. *Computers and Education: Artificial Intelligence*, 7, 100339. <https://doi.org/https://doi.org/10.1016/j.caeai.2024.100339>
- Yu, D. (2024). Research on the reform of university education pedagogy based on deep learning model in the context of information era. *Applied Mathematics and Nonlinear Sciences*, 9(1). <https://doi.org/10.2478/amns.2023.2.00765>
- Zhang, D.-W., Boey, M., Tan, Y. Y., & Jia, A. H. S. (2024). Evaluating large language models for criterion-based grading from agreement to consistency. *Npj Science of Learning*, 9(1), 79. <https://doi.org/10.1038/s41539-024-00291-1>