Journal of Inclusive Methodology and Technology in Learning and Teaching
ISSN 2785-5104
Anno 5 n.3 (2025)

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

When Artificial Intelligence Meets Metacognition: The Experimental Learning Unit Transforming Teaching and Evaluative Practices. The Maieutic Digital Toolkit

Maria Vittoria Isidori[1], Henry Muccini[1], Clara Evangelista[1], Donatella Visceglia[1]

[1] University of L'Aquila; mariavittoria.isidori@univaq.it; henry.muccini@univaq.it; clara.evangelista@guest.univaq.it; visceglia.donatella@gmail.com

***** Correspondence: mariavittoria.isidori@univaq.it

Abstract: The introduction of Artificial Intelligence (AI) into educational contexts necessitates careful instructional design, with a crucial focus on the teacher's metacognitive role and the urgent need to develop teacher Agency. This article describes the implementation of an Experimental Learning Unit (ELU) in a blended format for 320 teachers, designed to integrate AI tools with metacognitive reflection. The research adopted a two-phase approach: Phase 1 – Exploratory Try-Out (Pilot Study) on a small sample (N=30) for calibration and prototyping, followed by Phase 2 – Operational Experimentation on the extended sample. The focus is the Try-Out methodology and the preliminary results. Preliminary findings indicate that the Try-Out allowed for the optimization of module clarity and the usability of the AI tools. The operational experimentation shows a significant increase in teachers' professional metacognitive awareness (MAI-adaptive scale) and a more critical and reflective adoption of AI tools for assessment. The final output is the creation of an Operational Maieutic Digital Toolkit.

Keywords: Toolkit; AI and evaluation; MAI-adaptive; metacognition; teacher training

1. Introduction

The integration of Artificial Intelligence (AI) into educational contexts (AIED) is transforming pedagogies and demanding a profound reassessment of the teacher's role. While much academic discourse has rightfully focused on AI's potential to personalize student learning and optimize administrative tasks, the crucial area of how AI impacts the teacher's professional competence and reflective practice remains relatively underexplored (Labadze et All., 2023; Lu et All., 2024). Effective teaching in complex and inclusive classrooms requires more than technical skill; it necessitates professional metacognition—the teacher's ability to plan, monitor, and adapt their instructional actions reflectively (AlKanaan, 2022).

Current research often highlights two parallel challenges: first, the need for *AI literacy* (the knowledge of how AI works), and second, the practical application of AI *for improving teaching methods*. However, a significant gap persists in training models that merge these two necessities with the teacher's cognitive process. Teachers often

struggle to translate raw data from digital tools into meaningful instructional decisions, a challenge known as the *theory-practice gap* (Laupichler et All., 2022).

This work addresses this critical juncture by focusing on the theoretical axis of metacognition. We conceptualize AI not as a mere instrument, but as a *metacognitive mediator*. By providing objective data and actionable feedback on student performance or instructional design effectiveness, AI enhances the teacher's capacity to self-regulate and adapt their practices (Zhang et All., 2025; Isidori et All., 2024).

The primary objective of the project was to develop and validate a *training model capable of transforming tool-literacy* (the mere knowledge of tools) into *Teacher Design Agency*. Teacher Agency is defined here as the ability to act autonomously, intentionally, and critically within one's professional context, particularly when faced with new technologies (Alasgarova & Rzayev, 2025). This approach is significant because it shifts the focus from passively receiving AI *outputs* to actively using AI as an input for higher-order professional reflection.

This study describes the implementation of an *Experimental Learning Unit (ELU)* designed to integrate specific AI tools with metacognitive practices among a large cohort of in-service teachers. The research employed a rigorous *two-phase approach,* which included an initial *Exploratory Try-Out* for calibration, ensuring the ecological validity of the final model before the widespread *Operational Experimentation*.

The preliminary results affirm the effectiveness of the model. The Try-Out methodology proved essential in optimizing the usability of the AI tools and the clarity of the training modules. Subsequently, the operational experimentation demonstrated a *significant increase in teachers' professional metacognitive awareness* (measured using the MAI-adaptive scale) and a more critical and reflective adoption of AI tools for formative assessment. Ultimately, this work contributes a validated framework and an *Operational Digital Toolkit* that supports the sustained development of Teacher Design Agency in the digital era.

2. **Materials and Methods**

The following section provides a detailed description of the experimental design, the methodological phases and the instruments used. The purpose is to ensure sufficient detail for the potential replication and further development of the presented training model.

2.1. The Two-Phase Design and the Pilot Study

To achieve the primary goal of enhancing professional metacognition through the strategic integration of AI, an *Experimental Learning Unit (ELU)* was developed and implemented. The training was conducted over a total of 15 hours in a blended format (combining synchronous, in-person or video conference sessions with asynchronous activities). The overall study involved a total cohort of 320 in-service teachers.

The research design followed a rigorous, sequential methodological approach, articulated into two distinct phases to ensure the ecological validity and robustness of the final intervention model.

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

2.1.1. Phase 1: Exploratory Try-Out

The initial phase of the project was exclusively dedicated to the Exploratory Try-Out (Pilot Study). This phase utilized a reduced sample of N=30 teachers, who were selected based on their heterogeneity across different school levels and grades (Primary, Lower Secondary and Upper Secondary) to gather diversified feedback.

The main purpose of the Try-Out was not the collection of quantitative efficacy data, but rather the ecological validation and calibration of the ELU. Specifically, the aims were to: verify the clarity and time adherence of the theoretical content (Module 1: AI, ML, DL, NLP and Metacognition foundations).

Test the usability and instructional effectiveness of the specific AI tools proposed in Module 2 for self-regulation and instructional design support (*Maieutic Digital Toolkit*). Ensure that the instruments were comprehensible, culturally appropriate and free of ambiguity before wider administration. Table 1

Table 1. Composition of the Maieutic Digital Toolkit (AI & Generative Tools)

Category (Functional Focus)	*Example Tools (Free of Charge)*	*Key Pedagogical Purpose (Derived from Try-Out)*
Text and Ideation (LLM)	Google Gemini, Microsoft Co-pilot, Claude	Acts as a Metacognitive Partner for the teacher, supporting the design of instructional scaffolds and prompts
Search (Conversational Engines)	Perplexity AI, You.com	Supports Critical Research and contextualized analysis of sources, combating misinformation
Images and Multimedia (Generative AI)	Canva AI, Bing Image Creator	Enables the Rapid and Accessible Creation of visual and creative teaching resources for inclusivity.
Education (Applied AI)	Diffit, Curipod, Quizlet	Offers immediate support for Differentiation and data-informed Formative Assessment

2.1.2. Data Collection Methods

Qualitative data collection methods were prioritized during the Try-Out to facilitate rapid and iterative model refinement:

- Participant Observation: Direct monitoring during synchronous sessions was used to identify points of confusion, learning bottlenecks or practical difficulties encountered during the hands-on use of AI tools.

- Logbooks: Teachers were asked to document their experiences, subjective perceptions and specific difficulties encountered while using the AI tools, providing rich, continuous qualitative data.

- Qualitative Final Focus Group: A guided discussion was held across the three modules to gather in-depth feedback on the relevance, practical applicability, and perceived innovative charge of the ELU.

2.1.3. Phase 2: Operational Experimentation

Following the iterative optimization and calibration of the ELU based on the qualitative results of the Try-Out (as detailed in Section 3), the second phase involved implementation on the extended sample (N=290 teachers).

This phase utilized a pre-test/post-test with retesting design (T0, T1, T2) to measure effectiveness in terms of change in competencies, specifically concerning professional metacognition (as the primary dependent variable) and AI literacy (as a related competency).

3. Results

The presentation of results follows the sequential logic of the research design, starting with the crucial insights gained from the Pilot Study.

3.1. Qualitative Results of the Pilot Study (Try-Out)

Phase 1 (Try-Out) proved instrumental in ensuring the robustness and relevance of the training model. Key feedback led to necessary adjustments:

- Content Recalibration: Pilot teachers reported an excessive density of pure informatics concepts on AI. The ELU was recalibrated to shift the focus toward the pedagogical dimension of AI, reducing the emphasis on technical definitions.

- Tool Simplification: The difficulty in the autonomous use of AI tools was resolved by creating simplified operational sheets and transforming practical activities into guided exercises, ensuring AI was perceived as a supported resource rather than a source of complexity.

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

3.2. Quantitative Results of the Operational Experimentation

Preliminary results from Phase 2 (T1), realized thanks to the methodological solidity secured by the Try-Out, indicate a strong impact of the intervention on teachers:

- increase in Professional Metacognition: The pre-test/post-test comparison (T0 vs. T1) on the MAI-adaptive scale showed an average increase of 22% in scores related to the ability to Monitor and Regulate one's own instructional choices:

- reflective AI Adoption: Analysis of the post-test questionnaires (T1) showed that over 75% of teachers actively integrated AI tools into the design of formative assessments. Crucially, this adoption was not purely instrumental but occurred through the use of self-assessment rubrics and the integration of AI data with their own pedagogical expertise, aligning with the goal of developing Agency. Table.2

Table 2. AI Training Toolkit, Modules for Self-Regulation and Results.

Objective	- Evaluate the efficacy of a blended training program (10 hours) on AI-metacognition integration by measuring: 1. Teachers' metacognitive awareness (pre/post MAI-adaptive 2.0) 2. Operational competence in using AI tools for teaching.
Module 1: Foundations of Metacognition & AI	Contents: - Differences between AI-for-learning (e.g., data analysis with ML) vs. learning-about-AI (algorithmic literacy) - Examples: Gemini for metacognitive dialogue, Claude for study strategy feedback - Activity: ChatGPT-4 simulations for cognitive bias reflection, Google Gemini/Microsoft Copilot/Claude for ideation and text generation related to concepts.
Module 2: AI Tools for Self-Regulation	Contents: - Adaptive platforms (Diffit, Curipod, Quizlet) and planning tools (Notion.ai) - Evidence: 35% improvement with adaptive methods.

	- Activity: Creating study plans with AI, using Notion AI for self-assessment, Perplexity AI/You.com for research on self-regulation strategies.
Module 3: Metacognitive Evaluation with AI	Contents: - Predictive analysis of failure risk (Google's Gemini). - Auto-generated quizzes (Quizlet). - Activity: Analyzing student failure reports on learning styles (Gemini, AI), creating images/multimedia for feedback using Canva AI/Bing Image Creator.
Assessment -	Pre/Post: - MAI-adaptive 2.0 - AI-Integration questionnaire. - Post-Test: - MAI/re-test - Project work: Designing AI-enhanced lessons (Gemini 1.5) - Control group comparison.
Tools Used -	ChatGPT-4o, Gemini 1.5, Claude 2, Microsoft Copilot, Perplexity AI, You.com, Diffit, Curipod, Quizlet, Canva AI, Bing Image Creator (open-source models).
Key Metrics -	Improvement in MAI-adaptive 2.0 scores - Number of integrated AI tools (focus 2025-26 tools).
Expected Impact	- 25-30% increase in metacognitive awareness. - 22 AI tools (including Gemini) adopted by 75% of teachers.

4. Discussion

The preliminary findings from the Experimental Learning Unit (ELU) provide robust support for the central working hypothesis: that Artificial Intelligence, when integrated within a clear metacognitive framework, functions as an effective catalyst for professional reflection, rather than merely a technical utility. The significant increase in teachers' professional metacognitive awareness (measured by the MAI-adaptive scale) after the intervention directly addresses the need to bridge the theory-practice gap often cited in AIED literature (AlKanaan, 2022; Laupichler et All., 2022). Previous studies highlight that data literacy is a primary barrier for educators; our model overcame this by reframing AI-generated data not as a final answer, but as

a prompt for metacognitive inquiry (Alasgarova & Rzayev, 2025). The operational experimentation demonstrated that teachers could move beyond simply adopting tools (tool-literacy) to actively using the data to self-regulate their instructional design, thereby fostering Teacher Design Agency (Bhimdiwala et All., 2022; Dogan et All., 2025).

Crucially, the success of the extended experimentation was anchored by the Exploratory Try-Out (Pilot Study). The qualitative data collected in Phase 1 directly informed the recalibration of the ELU, ensuring the pedagogical focus and the ecological validity of the final model. This iterative approach is a key methodological implication, suggesting that complex training interventions must prioritize usability and applicability through pilot testing before full-scale deployment. By simplifying the Learning Analytics interface and contextualizing the AI feedback, the intervention reduced cognitive load and maximized the time available for higher-order professional reflection.

The most tangible implication of this research is the resulting Operational Digital Toolkit. This toolkit is not a mere repository of resources but represents the practical transposition of the research findings, designed to ensure the scalability and sustained impact of the training model.

The specific components of the toolkit reflect the key areas where AI integration proved most effective:

Simplified Operational Sheets: These tools ensure that ethical and pedagogical considerations are embedded directly into the how-to instructions for using generative and analytical AI tools, preventing purely instrumental adoption.

MAI-adaptive Templates: By providing practical examples of AI-integrated metacognitive assessment, the toolkit facilitates the shift from summative to formative assessment, aligning evaluation practices with the goal of student autonomy.

Design Reflection Matrix: This innovation operationalizes the metacognitive mediation of AI, compelling teachers to compare their intended learning design with the analytical feedback generated by the system. This actively supports the long-term development of Teacher Design Agency by making the teachers' professional choices explicit and measurable.

Future Research Directions. While the preliminary results at T1 are highly promising, the long-term sustainability of these metacognitive gains remains a key question. Therefore, future research steps will center on the analysis of the retesting data (T2), scheduled to validate the stability of the changes in metacognitive awareness over time. This will provide necessary insights into whether the training fosters deep, enduring shifts in professional practice or merely short-term behavioral adjustments. Furthermore, the findings lay the groundwork for formulating policy recommendations at the national level, advocating for integrating metacognitive training powered by AI into continuous professional development programs. Future studies should also explore the effectiveness of the Digital Toolkit across diverse national curricula and specific disciplinary contexts.

5. Conclusions

This study successfully developed and validated a two-phase training model that effectively integrates Artificial Intelligence and professional metacognition among

in-service teachers. The research demonstrates that AI functions as a crucial meta-cognitive mediator, significantly enhancing teachers' self-regulatory capabilities and fostering Design Agency. The rigorous Exploratory Try-Out ensured the practical robustness of the training, leading to a demonstrable increase in reflective AI adoption for evaluation. The resulting Operational Digital Toolkit provides a scalable framework to support the sustained transformation of teaching competencies in the digital era.

References

Alasgarova, R. & Rzayev, J. (2025). The implications of artificial intelligence for teacher agency and teacher-student relationships through the Technology Acceptance Model. *International Journal of Technology in Education and Science* (IJTES), 9(3), 450-473.

AlKanaan, H. M. N. (2022). Awareness regarding the implication of artificial intelligence in science education among pre-service science teachers. *International Journal of Instruction, 15*(3), 895–912.

Bhimdiwala, A., Neri, R. C., & Gomez, L. M. (2022). Advancing the design and implementation of artificial intelligence in education through continuous improvement. *International Journal of Artificial Intelligence in Education, 32*, 756–782.

Dogan, S., Nalbantoglu, U. Y., Celik, I., & Agacli Dogan, N. (2025). Artificial intelligence professional development: a systematic review of TPACK, designs, and effects for teacher learning. *Professional Development in Education*, *51*(3), 519–546.

Isidori, M. V., Muccini , H., Santelli, A. & Evangelista, C. (2024). Education and training for sustainability. Towards Artificial Intelligence: an exploratory investigation on teachers. *Form@re - Open Journal Per La Formazione in Rete*, 24(1), 294–300.

Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: systematic literature review. *International Journal of Educational Technology in Higher Education*, 20(1).

Laupichler, M. C., Aster, A., Schirch, J. and Raupach, T. (2022) 'Artificial intelligence literacy in higher and adult education: a scoping literature review', *Computers and Education: Artificial Intelligence*, 3, 100101.

Lu, J., Zheng, R., Gong, Z., & Xu, H. (2024). Supporting teachers' professional development with generative AI: The effects on higher order thinking and self-efficacy. *IEEE Transactions on Learning Technologies*, 17(3), 210–222.

Zhang, S., Ganapathy Prasad, P., & Schroeder, N. L. (2025). Learning About AI: A Systematic Review of Reviews on AI Literacy. *Journal of Educational Computing Research*, *63*(5), 1292-1322.