Giornale Italiano della Ricerca Educativa
Italian Journal of Educational Research

Generative Artificial Intelligence and Inclusive Co-design: A Study on Cognitive and Metacognitive Scaffolding in the Training of Special Education Teachers

Intelligenza artificiale generativa e co-progettazione inclusiva: uno studio sullo *scaffolding* cognitivo e metacognitivo nella formazione dei docenti di sostegno

Cinzia Referza
University of L'Aquila, (Italy)

OPEN ACCESS

Double blind peer review

Citation: Referza, C. (2025). Generative Artificial Intelligence and Inclusive Co-design: A Study on Cognitive and Metacognitive Scaffolding in the Training of Special Education Teachers. *Italian Journal of Educational Research*, 35, 177-192
https://doi.org/10.7346/sird-022025-p177

Received: July 14, 2025
Accepted: December 9, 2025
Published: December 27, 2025

Pensa MultiMedia / ISSN 2038-9744
https://doi10.7346/sird-022025-p177

Abstract
Generative Artificial Intelligence (GAI) is progressively reshaping education, opening up new scenarios for instructional and training innovation. Among its most promising applications, conversational tools such as ChatGPT show significant potential in supporting interdisciplinary instructional co-design, functioning as cognitive and metacognitive scaffolding for pre-service teachers. To bridge the gap between theory and practice in GAI-related training, an educational study was conducted involving 217 students enrolled in the Specialization Course for Teaching Support (9th Cycle, A.Y. 2023–2024, University of L'Aquila). The theoretical framework is grounded in socio-cultural constructivism, with reference to the Zone of Proximal Development (Vygotsky, 1978), the concept of scaffolding (Wood *et al.*, 1976), and the TPACK framework (Mishra & Koehler, 2006). A mixed methods approach was adopted, integrating quantitative and qualitative data collection and analysis (questionnaires, interviews, focus groups, and instructional artifacts developed with GAI support). Quantitative results indicate an increase in perceived self-efficacy regarding the use of digital technologies; qualitative findings highlight GAI's contribution to collaborative design and the development of inclusive teaching practices within interdisciplinary learning units. The discussion emphasizes the formative potential of GAI, as well as its criticalities, stressing the importance of a conscious, ethical, and methodologically sound use. The results, consistent with recent studies (Sperling *et al.*, 2024; Sadykova & Kayumova, 2024), offer valuable insights for integrating GAI into the TFA curriculum, helping to address a gap in empirical literature on inclusive instructional design supported by artificial intelligence.

Keywords: generative artificial intelligence, inclusive co-design, cognitive and metacognitive scaffolding, teacher education

Riassunto
L'intelligenza artificiale generativa (IAG) sta progressivamente ridefinendo l'istruzione, aprendo scenari inediti per l'innovazione didattica-formativa. Tra le applicazioni più promettenti, strumenti conversazionali come ChatGPT mostrano un potenziale significativo nel supportare la co-progettazione didattica interdisciplinare, fungendo da supporto cognitivo e metacognitivo (scaffolding) per i docenti in formazione. Per colmare il divario tra teoria e prassi nella formazione all'uso dell'IAG, è stata condotta una ricerca educativa su 217 corsisti iscritti al Corso di Specializzazione per le attività di sostegno didattico (IX Ciclo, a.a. 2023–2024, Università degli Studi dell'Aquila). Il quadro teorico si fonda sul costruttivismo socio-culturale, con riferimento alla zona di sviluppo prossimale (Vygotskij, 1978), al concetto di scaffolding (Wood *et al.*, 1976) e al framework TPACK (Mishra & Koehler, 2006). È stato adottato un approccio a metodi misti, con raccolta e analisi integrata di dati quantitativi e qualitativi (questionari, interviste, focus group e artefatti progettuali sviluppati con il supporto dell'IAG). I risultati quantitativi indicano un incremento dell'autoefficacia percepita rispetto all'uso delle tecnologie digitali; i dati qualitativi evidenziano il contributo dell'IAG alla progettazione collaborativa e allo sviluppo di pratiche didattiche inclusive all'interno delle unità di apprendimento interdisciplinari. La discussione sottolinea le potenzialità formative dell'IAG, ma anche le sue criticità, richiamando la necessità di un uso consapevole, etico e metodologicamente fondato. I risultati, coerenti con studi recenti (Sperling *et al.*, 2024, Sadykova & Kayumova, 2024), offrono spunti per integrare l'IAG nei curricula TFA, colmando una lacuna nella letteratura empirica sulla progettazione didattica inclusiva supportata dall'intelligenza artificiale.

Parole chiave: intelligenza artificiale generativa, co-progettazione inclusiva, *scaffolding* cognitivo e metacognitivo, formazione degli insegnanti

1. Introduction

The role of teachers has undergone a profound transformation in recent decades, reflecting the social, cultural, and technological shifts of contemporary society. Moving beyond the transmissive and authoritarian model of the past, today's teacher is increasingly recognized as a facilitator of learning, a designer of educational environments, and a cultural mediator, capable of integrating disciplinary knowledge, digital tools, and responses to special educational needs (Basilotta-Gómez-Pablos *et al.*, 2022; Laurillard, 2012; OECD, 2018; Schleicher, 2018).

This evolution reflects a significant epistemological shift that has progressively moved the focus from teaching to learning processes, adopting a student-centered perspective based on interaction, active knowledge construction, and the appreciation of diversity in school contexts (Biesta, 2010; Darling-Hammond *et al.*, 2017). Constructivist and socio-constructivist paradigms have encouraged the emergence of dialogic, reflective, and process-oriented conceptions of teaching (Hadar *et al.,* 2020), aligned with the growing attention to complexity and heterogeneity in classrooms (Florian & Black-Hawkins, 2011).

At the same time, the integration of digital technologies has reshaped traditional educational frameworks, redefining roles, relationships, and access to knowledge (Ertmer & Ottenbreit-Leftwich, 2010). Within this landscape, Generative Artificial Intelligence (GAI) is emerging as a promising resource capable of providing cognitive, metacognitive, and design-based support, especially in highly complex and heterogeneous learning environments—often described as "high-density inclusive settings" due to the coexistence of diverse educational needs and multi-level design challenges (Holmes *et al.*, 2019; Luckin & Holmes, 2016; Sánchez *et al.*, 2025; Tsirantonaki & Vlachou, 2025).

This study adopts the socio-cultural paradigm of learning as its theoretical framework, drawing on the concept of the Zone of Proximal Development (Vygotsky, 1978), understood as a dynamic space in which individual potential is activated through meaningful interactions with social mediators and cultural tools, and on the notion of scaffolding (Wood *et al.*, 1976), defined as temporary and adaptive support that facilitates the autonomous acquisition of complex skills. In this sense, GAI emerges as a dual scaffold—cognitive and metacognitive—capable of supporting inclusive instructional co-design in the training of future special education teachers.

GAI is conceived not only as a cognitive tool but also as a metacognitive facilitator, offering structured prompts, reflective cues, and feedback that can help pre-service teachers consciously regulate their thought processes, critically evaluate pedagogical choices, and collaboratively plan inclusive practices. This scaffolding function contributes to the formation of a professional identity that is reflective, autonomous, and sensitive to the complexity of inclusive education.

From this perspective, the study aims to explore the potential of GAI as a form of cognitive and metacognitive scaffolding in instructional design, with a specific focus on three dimensions: pedagogical reflection, cross-disciplinary integration, and the adoption of inclusive practices. The methodological design follows a mixed methods approach, integrating quantitative and qualitative data (questionnaires, interviews, observations, focus groups, and analysis of digital artifacts).

The contribution seeks to fill a gap in the scientific literature, which still lacks empirical evidence on the use of GAI in initial teacher education—particularly concerning inclusive, interdisciplinary, and reflective instructional design. The expected results aim to provide operational guidelines for updating the TFA curricula, promoting a critical, ethical, and pedagogically grounded use of emerging technologies in education (Selwyn, 2019; Williamson & Eynon, 2020; Eickelmann *et al.*, 2019).

2. Theoretical Framework

In recent years, Generative Artificial Intelligence (GAI) has attracted growing interest in the field of education as a technology capable of expanding the possibilities for designing, personalizing, and assessing teaching and learning processes (Zawacki-Richter *et al.*, 2019; Popenici & Kerr, 2017; Cabero-Almenara *et al.,* 2023). Unlike traditional systems based on predictive logic, GAI generates content autonomously in natural language and in a context-sensitive manner, stimulating higher-order thinking processes such as critical reflection, divergent thinking, and collaborative revision (Luckin & Holmes, 2016; Kovanovi *et al.*, 2023; Rashid *et al.*, 2024). These features make it particularly suitable for supporting the active construction of knowledge, the development of pedagogical thinking, and the activation of metacognitive strategies such as planning, monitoring, and self-assessment (Holmes *et al.*, 2019; Levin *et al.* (2025); Hadar *et al.*, 2020).

This contribution is grounded in the socio-cultural paradigm of learning, with reference to Vygotsky's (1978) concept of the Zone of Proximal Development (ZPD), conceived as a dynamic space in which learners' potential is actualized through mediated interactions with more competent others and cultural tools. Within this framework, the concept of scaffolding, introduced by Wood, Bruner, and Ross (1976), becomes particularly relevant. Scaffolding refers to the temporary and adaptive support provided to learners to help them achieve tasks and develop competencies they could not master independently.

In digital learning environments, scaffolding can take both cognitive and metacognitive forms. Cognitive scaffolding includes support for conceptual reorganization, planning, and problem-solving; metacognitive scaffolding, on the other hand, supports processes such as critical reflection, self-assessment, and the regulation of thinking (Hammond & Gibbons, 2005; Hmelo-Silver, Duncan, & Chinn, 2007; Abdelghani *et al.*, 2023**;** Kovanovi *et al.*, 2023). In this light, GAI can function as both a cognitive and metacognitive scaffold, offering prompts, guiding questions, and structured feedback aligned with the learner's level of development (Reich & Ito, 2017).

Another relevant theoretical lens is the TPACK model (Technological Pedagogical and Content Knowledge), which underscores the need for the integrated development of disciplinary, pedagogical, and technological knowledge in teacher education (Mishra & Koehler, 2006; Koehler *et al.,* 2013; Gatete, 2025). Within this model, GAI can enhance the reflective capacity of pre-service teachers, fostering the integration of content, method, and technology through dialogic interaction and generative feedback (Gatete, 2025).

Furthermore, GAI-supported instructional co-design is understood as an authentic context for professional learning that fosters teacher agency, peer collaboration, and metacognitive competencies essential for intentional and inclusive instructional planning (Kafai & Resnick, 1996; Kirschner & Erkens, 2013; Johnson & Johnson, 2009). This perspective aligns with the concept of "pedagogical mindfulness" in technology integration, as discussed by Laurillard (2012), where digital tools are evaluated not only for their efficiency but also for their ability to support thoughtful pedagogical decision-making.

In inclusive settings, GAI has also been recognized for its ability to personalize learning pathways, differentiate materials, and generate accessible content based on students' profiles and educational needs (Luckin & Holmes, 2016; Holmes *et al.*, 2019).

In assessment contexts, it can support the creation of authentic tasks, timely feedback, and self-reflective activities that nurture students' metacognitive awareness (Whitelock & Cross, 2012; Mahamuni *et al.*, 2024). At the same time, critical issues remain regarding algorithmic transparency, ethical use, and digital inequalities (Binns *et al.,* 2018; Selwyn, 2019; Alon-Barkat & Busuioc, 2023).

Williamson & Eynon, 2020). These challenges reinforce the need for teacher education to go beyond technical competence and foster ethical awareness, critical thinking, and conscious regulation of mental processes (Livingstone & Sefton-Green, 2016).

In conclusion, GAI is not merely a technological tool but a potential mediator of cognitive and metacognitive development, contributing to the emergence of a reflective, autonomous, and ethically grounded professional identity in future teachers (Eickelmann *et al.,* 2019).

Its capacity to function as both cognitive and metacognitive scaffolding positions GAI as a powerful lever in teacher education, particularly in inclusive instructional design contexts. By supporting processes of reflection, goal setting, and adaptive decision-making, GAI enhances not only the quality of instructional planning but also the professional agency of pre-service teachers.

Although empirical research on GAI in education is still in its early stages, especially within initial teacher training, this contribution aims to fill part of this gap by critically analyzing the pedagogical and reflective value of GAI in interdisciplinary co-design processes aimed at inclusion. This requires a systematic and ethically conscious approach to GAI integration, one that prioritizes its educational potential over its technical novelty.

3. Research Context and Sample Characteristics

The study was conducted at the University of L'Aquila and involved 217 pre-service teachers enrolled in the Specialization Course for Teaching Support in lower and upper secondary education (9th Cycle, A.Y. 2023–2024). Among the participants, 77.9% identified as female and 22.1% as male. The average age was 36 years (SD = 7.4), with a range between 24 and 57 years.

All participants held a master's degree or an equivalent qualification, in line with the admission requirements for the TFA program. Additionally, 4.1% had completed a Ph.D., indicating the presence—albeit limited—of highly qualified academic profiles.

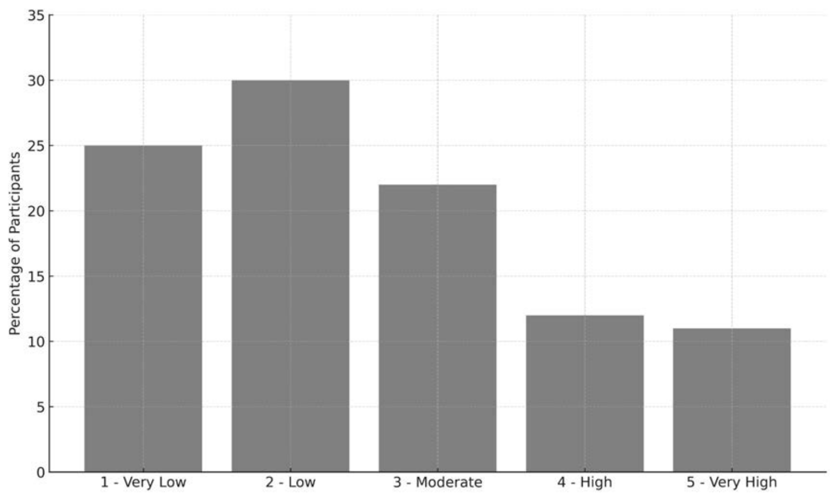
Regarding prior training in educational technologies, 32.7% had attended at least one university course or professional development activity focused on the pedagogical use of digital tools. However, only 2.8% had participated in training specifically dedicated to the use of artificial intelligence (AI) in education.

The majority of the sample (67.2%) had not taken part in structured training on the pedagogical integration of technologies, confirming what is already reported in the literature about the limited diffusion of pedagogical-digital competencies related to AI in initial teacher education (Holmes *et al.,* 2019; Zawacki-Richter *et al.,* 2019; Hwang & Tu, 2021).

With respect to prior professional experience, 59.4% of the trainees reported having previously taught in schools, while the remaining 40.6% were at the beginning of their teaching careers. This heterogeneity is significant, as previous experiences influence both the approach to technology and the pedagogical representations associated with its use (Instefjord & Munthe, 2017; Tondeur *et al.,* 2019; Greenhow & Askari, 2017).

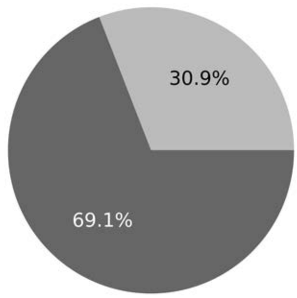
As for familiarity with AI, participants self-reported relatively low levels of knowledge: more than 55% rated their understanding of concepts such as machine learning, natural language processing, and generative systems between 1 and 2 on a 5-point Likert scale. This finding confirms the existence of a cognitive gap already identified in preious studies (Manca & Ranieri, 2017; Vuorikari *et al.,* 2022).

This pattern of limited AI literacy among participants is visually represented in Graf. 1**,** which illustrates the distribution of self-assessed familiarity with core artificial intelligence concepts.



Graf. 1: Self-Reported Familiarity with Core AI Concepts

Nevertheless, interest in gaining a critical understanding of AI in education was high: 67.8% expressed an interest level of 4 or 5 out of 5, indicating strong motivation to deepen their knowledge of a topic perceived as innovative, urgent, and strategically relevant to their professional development (Abdelghani *et al.*, 2023; Tsirantonaki & Vlachou, 2025; Kimmons *et al.,* 2023). The proportion of participants who had already experimented with generative AI tools in educational contexts is presented in Graf. 2, highlighting the current extent of practical engagement.



Graf. 2: Use of Generative AI Tools in Educational Contexts

Additionally, 30.9% of the participants had already experimented with generative AI tools—such as ChatGPT, Canva Magic Write, or Grammarly—in educational contexts. The remaining 69.1% reported never having used them for teaching purposes, highlighting a still underexplored training potential. This finding underscores the need for a structured training intervention aimed at developing critical, metacognitive, operational, and ethically informed competencies for the use of GAI in instructional design (Luckin & Holmes, 2016; Holmes *et al.,* 2019).

4. Methodological Approach and Mixed Methods Design

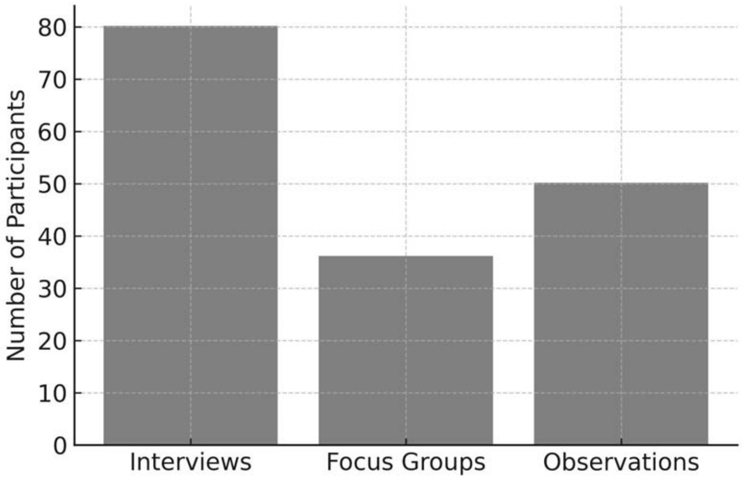
This study adopts a mixed methods research design (Creswell & Plano Clark, 2018), considered particularly suitable for investigating complex and multidimensional educational phenomena. This approach enables the integration of quantitative and qualitative data according to a logic of methodological triangulation and epistemological complementarity (Johnson & Onwuegbuzie, 2004; Teddlie & Tashakkori, 2009), responding to the growing demand for research that combines analytical rigor with contextual depth (Miles *et al*., 2014; Plano Clark & Ivankova, 2016).The study employed a convergent parallel design, in which numerical and narrative data were collected simultaneously, analyzed independently, and then integrated during the interpretation phase (Fetters *et al.,* 2013). This configuration allowed for the appreciation of the transformative and experiential dimensions of teacher training, while maintaining the empirical consistency needed to identify patterns and trends across data sets (Onwuegbuzie & Leech, 2006; Greene, 2007).

For the quantitative component, an online questionnaire was administered to the 222 students enrolled in the Specialization Course for Teaching Support (9th Cycle, A.Y. 2023–2024), with a response rate of 96.9% (n = 217). The questionnaire was developed based on a systematic review of international literature and validated measurement scales assessing technological self-efficacy (Bandura, 2006), pedagogical integration of digital tools (Ertmer & Ottenbreit-Leftwich, 2010), and attitudes toward digital teaching (Scherer *et al.,* 2019; Rubach & Lazarides, 2021).

The items explored the participants' educational background, previous experiences with educational technologies, familiarity with key AI concepts, perceived self-efficacy in AI-supported instructional design, and the activation of metacognitive strategies. Data were collected using a five-point Likert scale (1 = not at all to 5 = very much), in line with existing research on teacher beliefs and practices related to educational technologies (Tondeur *et al.,*2012; Røkenes & Krumsvik, 2016).

The qualitative component included three data collection methods: semi-structured interviews (n = 80), small-group focus groups (n = 36), and systematic classroom observations (n = 50).

The distribution of participants involved in the three qualitative data collection modalities is illustrated in Graf. 3, offering a visual overview of the sample structure across interviews, focus groups, and classroom observations.



Graf. 3: Participants in qualitative data collection methods

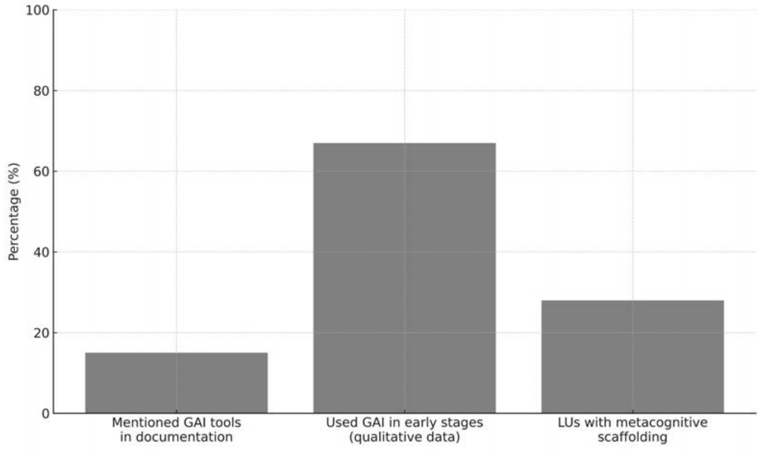
These approaches allowed for triangulation and the exploration of subjective meaning-making and reflective processes. Interview participants were selected using purposeful sampling (Patton, 2015), and the number was determined based on theoretical saturation (Guest *et al.,* 2006).

Focus groups (6 participants each, with one group of 5) facilitated the negotiation of shared interpretations and values (Krueger & Casey, 2014; Barbour, 2014), while classroom observations were guided by a structured rubric inspired by Derry *et al.* (2010) and Cohen *et al.* (2018).Particular attention was paid to the presence of scaffolding mechanisms, including peer modeling, generative prompts, and structured GAI-supported reflection. These mechanisms were analyzed through a framework informed by socio-constructivist perspectives on mediated learning (Hammond & Gibbons, 2005; Mercer, 2013).

As part of the documentary analysis, 50 interdisciplinary Learning Units (LUs) were reviewed, each created by groups of 4–5 participants. Evaluation was based on a multidimensional rubric drawing from four widely recognized frameworks: the TPACK model (Mishra & Koehler, 2006), Universal Design for Learning (CAST, 2018), DigCompEdu (European Commission, 2017), and the Italian Guidelines for School Inclusion (MIUR, 2017). Six dimensions were analyzed: pedagogical coherence, inclusivity, interdisciplinarity, methodological innovation, GAI integration, and the inclusion of metacognitive scaffolding prompts. A hybrid content analysis approach was applied, combining deductive and inductive coding (Mayring, 2014; Saldaña, 2021), and aligning with recommendations for qualitative rigor in educational research (Nowell *et al.,* 2017).

The convergence between quantitative and qualitative results, the variety of instruments employed, and the scale of the sample represent core strengths of the study's design (Guba & Lincoln, 1989; Tracy, 2010). Quality assurance for the qualitative data was ensured through source triangulation, inter-coder reliability checks, and member checking (Lincoln & Guba, 1985).Cronbach's alpha confirmed the internal consistency of the self-efficacy scale (= 0.89). Statistical analyses revealed a significant improvement in digital self-efficacy from the initial phase ($M = 2.61$; $SD = 0.87$) to the final phase ($M = 3.94$; $SD = 0.76$), $p < 0.001$ (paired sample t-test). The LUs scored on average: 4.3 for inclusiveness, 4.2 for pedagogical coherence, 4.1 for interdisciplinarity, 3.9 for innovation, and 3.8 for GAI integration.

Notably, 28% of the LUs included explicit metacognitive scaffolding strategies such as checklists, self-assessment tools, and reflective questions supported by GAI . The frequency of GAI-related components emerging from the analysis of Learning Units is summarized in Graf.4**,** highlighting the integration of generative tools and metacognitive strategies across instructional artifacts.



Graf. 4: GAI-Related Elements in Learning Units

Finally, the ICT module was conducted entirely by a single instructor, ensuring methodological continuity and coherence in the integration of metacognitive prompts and scaffolding mechanisms across all laboratory phases.

5. Integrated Descriptive Data Analysis

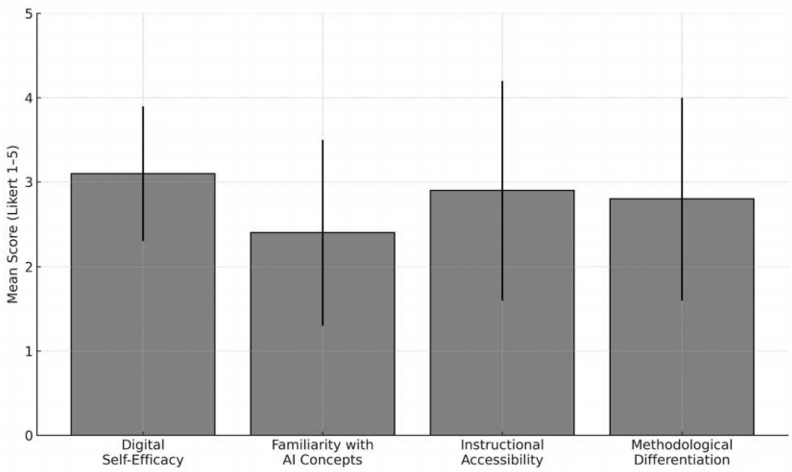
5.1 Quantitative Data Analysis – Pre-intervention

Data collected from the pre-intervention questionnaire administered to 217 pre-service teachers were analyzed using univariate descriptive statistics.

Ordinal variables (measured on a 5-point Likert scale) were examined through arithmetic mean, standard deviation, median, and mode; for categorical variables, absolute and relative frequencies were reported.

Participants showed moderate perceived self-efficacy in using digital technologies ($M = 3.1$; $SD = 0.8$), and low familiarity with AI concepts ($M = 2.4$; $SD = 1.1$), with the distribution skewed towards lower values (skewness = 0.87).

Dimensions such as instructional accessibility ($M = 2.9$; $SD = 1.3$) and methodological differentiation ($M = 2.8$; $SD = 1.2$) revealed considerable variance, indicating a heterogeneous background among participants. The Shapiro–Wilk test confirmed normality ($p > .05$), affirming the statistical robustness of the mean values (DeCarlo, 1997). These trends in self-efficacy and AI familiarity are visually summarized in Graf. 5, which illustrates participants' baseline perceptions across the key instructional dimensions.



Graf. 5: Digital Self-Efficacy and AI-Related Dimensions

These findings reflect the need for structured, reflective training interventions aimed at fostering critical digital competence and pedagogical inclusivity (Buckingham, 2015; Rachbauer *et al.*, 2025).

5.2 Qualitative Data Analysis: Metacognitive Processes and Participants' Representations

The qualitative corpus included 80 semi-structured interviews, 36 focus groups, and 50 classroom observations. An inductive thematic analysis was applied (Braun & Clarke, 2006), with double-blind coding carried out by independent researchers. Intercoder agreement was high (Cohen's = 0.81), indicating reliable interpretation (McHugh, 2012; Saldaña, 2021). Five recurring themes were identified:

Openness to innovation: Participants expressed enthusiasm for GAI, seeing it as a creative and engaging educational tool (Popenici & Kerr, 2017).

Tension between control and creativity: While appreciating GAI's support in content generation, participants expressed concern over diminished epistemological control (Selwyn, 2019; Williamson & Eynon, 2020).

Need for pedagogical frameworks: Participants called for methodological guidance in using GAI ethically and effectively, confirming the relevance of the TPACK model (Mishra & Koehler, 2006; Koehler *et al.,* 2013).

Co-design as metacognitive scaffolding: Collaborative lesson planning promoted self-regulation and professional reflection, supported by GAI suggestions (Schön, 1983; Laurillard, 2012; Mezirow, 1991).

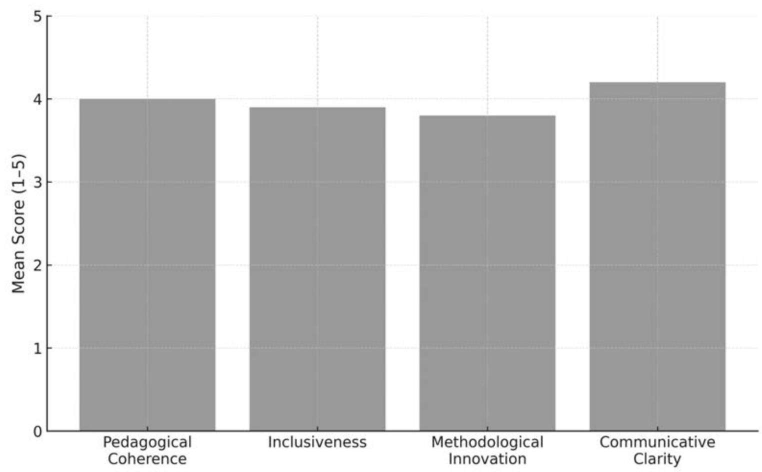
Ethical concerns: Interviewees highlighted the need to address privacy, source reliability, and bias, in line with recent literature (Binns *et al.,* 2018; Alon-Barkat & Busuioc, 2023; Floridi & Cowls, 2019).

5.3 Analysis of Instructional Artifacts: Learning Units

Fifty interdisciplinary Learning Units (LUs) were analyzed using a 5-point rubric assessing four dimensions: pedagogical coherence, inclusiveness, methodological innovation, and communicative clarity.

Average scores ranged from 3.8 to 4.2, with highest values in communicative clarity (M = 4.2; SD = 0.6) and greater dispersion in inclusiveness (M = 3.9; SD = 0.7), suggesting uneven familiarity with Universal Design for Learning (CAST, 2018). Only 15% of LUs explicitly mentioned GAI tools, but 67% of interviews and 66% of observations reported their use in brainstorming, planning, or editing.

Notably, 28% of the LUs included explicit metacognitive scaffolding strategies such as structured reflection prompts (e.g., 'What makes this inclusive?'), planning templates, or self-monitoring tools generated with GAI support (Wood *et al.*, 1976; Kovanovi *et al.,* 2023). The evaluation outcomes of the Learning Units (LUs) are presented in Graf. 6, highlighting differential performance across core pedagogical dimensions.



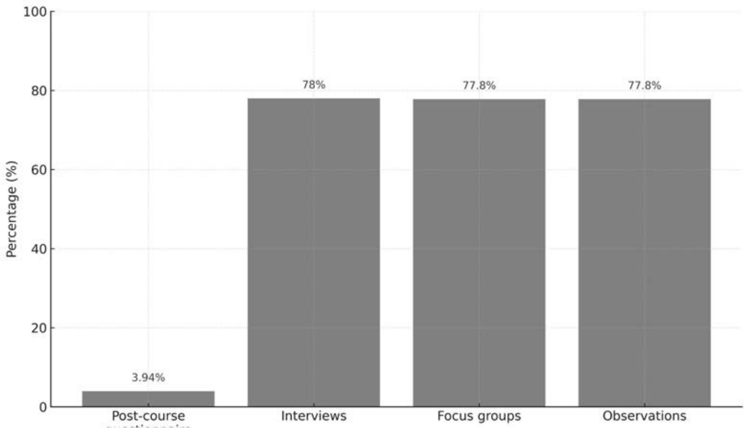
Graf.6: Average evaluation scores for Learning Units across four criteria (N = 47)

5.4 Results Integration

The triangulation of quantitative, qualitative, and artifact-based data revealed a strong convergence in findings.

High perceived digital self-efficacy reported in the post-course questionnaire (M = 3.94; SD = 0.76) was consistent with qualitative insights from 78% of interviews, 77.8% of focus groups, and 77.8% of observations, where participants described improved confidence in digital lesson planning. This consistency enhances the study's internal validity and transferability (Guba & Lincoln, 1989; Tracy, 2010), and highlights GAI's potential to serve as dynamic scaffolding for cognitive and metacognitive development (Luckin & Holmes, 2016; Tsirantonaki & Vlachou, 2025; Eickelmann *et al.,* 2019).

The findings support the claim that GAI, when used intentionally, contributes to inclusive instructional innovation and reflective professional identity formation. This convergence across data sources—reflected in the alignment between quantitative and qualitative findings —illustrates the perceived growth in digital self-efficacy and the transformative impact of GAI integration in instructional planning, as illustrated in Graf. 7.



Graf. 7: Convergence of Quantitative and Qualitative Evidence

6. Results and Discussion

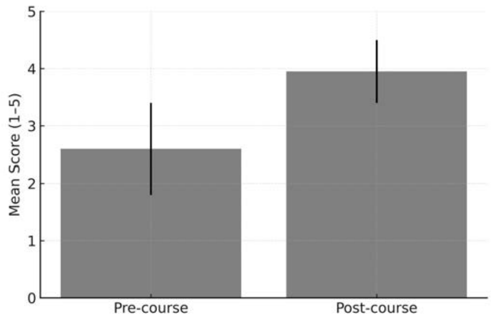
6.1 Quantitative analysis: perceptions and digital self-efficacy

The pre-intervention questionnaire (N = 217) revealed moderate levels of perceived self-efficacy in using digital technologies (M = 3.1; SD = 0.8), with lower confidence reported in the areas of instructional personalization through AI (M = 2.8; SD = 1.1) and digital accessibility (M = 2.9; SD = 1.3). This highlights a superficial and tool-oriented approach to educational technologies (Alon-Barkat & Busuioc, 2023; Vuorikari *et al.,* 2022).

As shown in Graf. 8, the shift in self-perceived digital competence before and after the course illustrates the effectiveness of the training intervention in strengthening participants' confidence in AI-supported instructional design.

The skewness coefficient (0.87) pointed to a concentration of responses toward lower values, suggesting limited digital readiness among a portion of the cohort. The Shapiro–Wilk test (p > .05) confirmed normality in data distribution, supporting the reliability of mean comparisons (DeCarlo, 1997). These results align with research noting the persistent gap in pre-service teacher training regarding educational AI (Zawacki-Richter *et al.*, 2019; Holmes *et al.,* 2019; Cabero-Almenara *et al.,* 2023).

Given the increasing centrality of AI in educational contexts, these findings underscore the need for structured and critically oriented professional development programs aimed at fostering reflective and metacognitive digital competence (Buckingham, 2015; Rachbauer *et al.*, 2025; Eickelmann *et al.*, 2019).



Graf. 8:Pre–Post Digital Self-Efficacy

6.2 Qualitative Analysis: Metacognitive Processes and Representations

Thematic analysis (Braun & Clarke, 2006) of 80 semi-structured interviews, one large-scale focus group (n = 36), and 50 observation reports identified five recurrent categories:

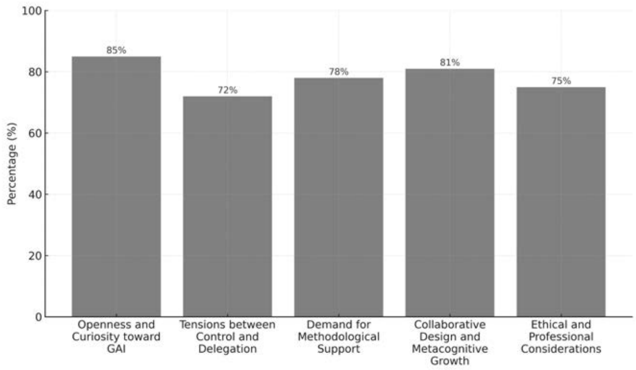
a) Openness and Curiosity toward GAI
Participants highlighted the novelty and potential of GAI in fostering creativity, personalization, and student engagement—echoing prior findings on technology-driven motivation (Ertmer & Ottenbreit-Leftwich, 2010).
b) Tensions between Control and Delegation
A recurring tension emerged between leveraging AI-generated content and the perceived loss of epistemic authority, resonating with critiques on automation and decision-making in education (Selwyn, 2019; Williamson & Eynon, 2020; Prensky, 2023).

c) Demand for Methodological Support
Participants consistently requested clearer pedagogical frameworks for the intentional integration of GAI, reflecting the necessity of structured guidance aligned with the TPACK model (Mishra & Koehler, 2006; Koehler *et al.,* 2013; Gatete, 2025).

d) Collaborative Design and Metacognitive Growth
GAI-supported instructional co-design was perceived as a catalyst for professional growth, enabling participants to engage in reflective planning, strategic thinking, and critical evaluation—central components of metacognitive regulation (Hammond & Gibbons, 2005; Mezirow, 1991; Sánchez *et al.*, 2025).

e) Ethical and Professional Considerations
The ethical use of GAI emerged as a cross-cutting concern, with participants raising issues related to authorship, bias, and data privacy, mirroring recent scholarly discourse on AI ethics in education (Binns *et al.,* 2018; Alon-Barkat & Busuioc, 2023).

The recurrence of each of these themes across data sources is illustrated in Graf. 9, offering a synthesized view of the qualitative trends that emerged from the analysis.

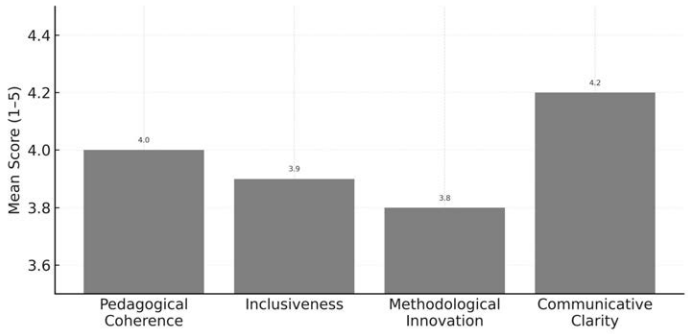


Graf. 9: Percentage of Participants Referencing Thematic Categories

6.3 Instructional Artifacts: Integrating GAI into Learning Units

The 47 Learning Units (LUs) were analyzed through a multidimensional rubric (scale 1–5), evaluating pedagogical coherence, inclusiveness, innovation, and communicative clarity. The average scores ranged from 3.8 to 4.2, with the highest mean for communicative clarity (M = 4.2; SD = 0.6), while inclusiveness (M = 3.9; SD = 0.7) showed greater dispersion, suggesting varied implementation of Universal Design for Learning (CAST, 2018; Hsu *et al.*, 2023).

As illustrated in Graf. 10, the comparative overview of the evaluation criteria applied to the Learning Units underscores key areas of pedagogical strength and highlights opportunities for further innovation.



Graf. 10 : Mean Scores of Learning Units by Evaluation Criteria

These artifacts support the dual cognitive and metacognitive scaffolding role of GAI, especially in inclusive co-design contexts (Luckin & Holmes, 2016; Kirschner & Erkens, 2013).

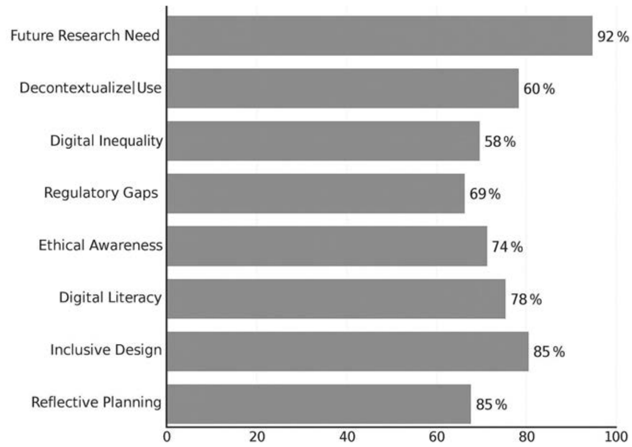
6.4 Integrated Discussion and Educational Implications

The consistency across quantitative, qualitative, and artifact-based data confirms the robustness of the mixed methods design (Greene *et al.*, 1989; Denzin, 1978). The integrated findings suggest that GAI can serve as a pedagogical catalyst—enhancing reflection, planning, and instructional inclusion—if used critically and ethically (Livingstone & Sefton-Green, 2016; Holmes *et al.,* 2019).

The integration of dedicated modules on AI in TFA curricula is strongly recommended, not only to bridge existing digital literacy gaps, but also to foster an ethically informed, pedagogically grounded, and metacognitively aware teaching profession (Eickelmann *et al.,* 2019; Sánchez *et al.*, 2025).

However, persistent challenges remain, including digital inequalities, regulatory uncertainty, and risks of decontextualized AI use (Alon-Barkat & Busuioc, 2023). The cross-sectional nature of this study also limits conclusions about long-term efficacy.

Future studies should explore the role of GAI in real-world classrooms, focusing on its contribution to student engagement, teacher agency, and inclusive planning over time (OECD, 2021; Holmes *et al.*, 2019). A synthesis of the primary pedagogical implications and open challenges emerging from the study is presented in Graf.11, reflecting the integrated insights of the mixed-methods analysis.



Graf. 11: Educational implications and challenges of integrating GAI into teaching

7. Conclusions

The findings of this study confirm the educational potential of Generative Artificial Intelligence (GAI) as a multifunctional pedagogical tool, particularly in enhancing inclusive, interdisciplinary, and reflective instructional design. When intentionally and critically integrated into teacher education programs, GAI can function as both cognitive and metacognitive scaffolding, supporting processes such as planning, conceptual reorganization, reflective evaluation, and decision-making (Wood *et al.,* 1976; Hammond & Gibbons, 2005; Kovanovi *et al.,* 2023).

Specifically, the analysis revealed GAI's ability to activate essential metacognitive functions—such as conscious goal-setting, monitoring of cognitive processes, and critical self-assessment of pedagogical choices—contributing to the development of a reflective, adaptive, and professionally autonomous teacher identity (Mezirow, 1991; Laurillard, 2012; Sánchez *et al.*, 2025).

These findings reinforce the strategic value of incorporating specific modules on AI in education within TFA (Specialization Courses for Support Teaching) curricula. Such training should not be limited to technical aspects but should instead promote a critical, ethically grounded, and pedagogically informed use of GAI (Selwyn, 2019; Eickelmann *et al.*, 2019). Training programs should explicitly include activities aimed at promoting metacognitive regulation, such as structured reflection prompts, iterative planning cycles, and collaborative design tasks facilitated by generative tools (Gatete, 2025; Kirschner & Erkens, 2013).

The study also highlights the dual role of GAI as a dialogic and adaptive scaffold—simultaneously supporting content generation and professional reflection. In this capacity, GAI contributes to the construction of inclusive learning environments while strengthening reflective teacher agency (Luckin & Holmes, 2016; Holmes *et al.,* 2019).

Nevertheless, important challenges remain: uneven digital literacy among pre-service teachers, the lack of an updated regulatory and ethical framework, and the potential risk of instrumental or uncritical use of GAI (Alon-Barkat & Busuioc, 2023). Furthermore, the absence of a longitudinal follow-up limits our understanding of the sustained impact of GAI on instructional practices and professional identity development.

Future research should address these gaps by exploring GAI's long-term influence in real-world educational contexts, with a specific focus on student motivation, inclusive planning, and reflective teaching practices. Particular attention should be given to the evolution of scaffolding mechanisms over time, to determine how GAI can support the dynamic and complex nature of contemporary teaching with critical awareness and pedagogical intentionality (Tsirantonaki & Vlachou, 2025; OECD, 2021; Holmes *et al.*, 2019).

Bibliografia

Abdelghani, R., Sauzéon, H., & Oudeyer, P.-Y. (2023*).* Generative AI in the classroom: Can students remain active learners? *arXiv preprint arXiv:2310.03192*. https://arxiv.org/abs/2310.03192

Alon-Barkat, S., & Busuioc, M. (2023). Human–AI interactions in public sector decision making: "Automation bias" and "Selective adherence" to algorithmic advice. *Journal of Public Administration Research and Theory*, *33*(1), 153–169.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 5, pp. 307–337). Greenwich, CT: Information Age Publishing.

Barbour, R. (2014). *Introducing qualitative research: A student's guide* (2nd ed.). London: SAGE.

Basilotta-Gómez-Pablos, V., Matarraz, M., Casado-Aranda, L.-A., & Otto, A. (2022). Teachers' digital competencies in higher education: A systematic literature review. *International Journal of Educational Technology in Higher Education, 19*(8).

Biesta, G. J. J. (2010). *Good education in an age of measurement: Ethics, politics, democracy.* Boulder, CO: Paradigm Publishers.

Binns, R., Veale, M., Van Kleek, M., & Shadbolt, N. (2018). 'It's reducing a human being to a percentage': Perceptions of justice in algorithmic decisions. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). New York: ACM Press.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.

Buckingham, D. (2015). *The media education manifesto.* Cambridge: Polity Press.

Cabero-Almenara, J., Gutiérrez-Castillo, J. J., Barroso-Osuna, J., & Rodríguez-Palacios, A. (2023). Digital teaching competence according to the DigCompEdu framework: comparative study in different Latin American universities. *Journal of New Approaches in Educational Research*, *12*(2), 276-291.

CAST. (2018). *Universal Design for Learning Guidelines version 2.2.* Wakefield, MA: CAST.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). London: Routledge.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Thousand Oaks, CA: SAGE.

Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2017). Implications for educational practice of the science of learning and development. *Applied Developmental Science*, *24*(2), 97–140.

DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Methods*, 2(3), 292–307.

Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods* (2nd ed.). New York: McGraw-Hill.

Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., Hall, R., Koschmann, T., Lemke, J. L., Sherin, M. G., & Sherin, B. L. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *Journal of the Learning Sciences*, *19*(1), 3–53.

Eickelmann, B., Bos, W., Gerick, J., Goldhammer, F., Schaumburg, H., Schwippert, K., Senkbeil, M., & Vahrenhold, J. (Eds.). (2019). *ICILS 2018 #Deutschland. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im zweiten internationalen Vergleich und Kompetenzen im Bereich Computational Thinking.* Münster: Waxmann.

Ertmer, P. A., & Ottenbreit-Leftwich, A. T. (2010). Teacher technology change: How knowledge, confidence, beliefs, and culture intersect. *Journal of Research on Technology in Education*, *42*(3), 255–284.

European Commission. (2017). *European Framework for the Digital Competence of Educators: DigCompEdu.* Luxembourg: Publications Office of the European Union.

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs: Principles and practices. *Annals of Family Medicine*, *11*(6), 574–581.

Florian, L., & Black-Hawkins, K. (2011). Exploring inclusive pedagogy. *Cambridge Journal of Education*, *41*(2), 147–162.

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, *1*(1).

Gatete, O. (2025). Revisiting TPACK: A critical review and contextual extension for the digital age. *Journal of Educational Technology.* Advance online publication.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, *11*(3), 255–274.
Greene, J. C. (2007). *Mixed methods in social inquiry.* San Francisco, CA: Jossey-Bass.
Greenhow, C., & Askari, E**.** (2017). Learning and teaching with social network sites: A decade of research in K–12 related education. *Education and Information Technologies, 22*(2), 623–645.
Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation.* Newbury Park, CA: Sage.
Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? *Field Methods*, *18*(1), 59–82.
Hadar, L. L., Ergas, O., Alpert, B., & Ariav, T. (2020). Rethinking teacher education in a VUCA world: Student teachers' social-emotional competencies during the Covid-19 crisis. *European Journal of Teacher Education*, *43*(4), 573-586.
Hammond, J., & Gibbons, P. (2005). Putting scaffolding to work: The contribution of scaffolding in articulating ESL education. *Prospect*: *An Australian Journal of TESOL*, *20*(1), 6–30.
Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning. *Educational Psychologist*, 42(2), 99–107.
Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Boston, MA: Center for Curriculum Redesign.
Hsu, Y.-C., & Ching, Y.-H. (2023). Generative Artificial Intelligence in Education, part one: The dynamic frontier. *TechTrends, 67*(1).
Hwang, G.-J., & Tu, Y.-F. (2021). Roles and research trends of artificial intelligence in education: A bibliometric mapping analysis and systematic review. *Computers & Education Artificial Intelligence, 2*, 100012.
Instefjord, E., & Munthe, E. (2017). Educating digitally competent teachers: A study of integration of professional digital competence in teacher education. *Teaching and Teacher Education*, *67*, 37–45.
Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher*, *38*(5), 365–379.
Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, *33*(7), 14–26.
Kafai, Y. B., & Resnick, M. (1996). *Constructionism in practice. Designing, thinking, and learning in a digital world.* Mahwah, NJ: Routledge.
Kimmons, R., Rosenberg, J. M., & Allman, B. A. (2023). Artificial intelligence and education: A review of research in the 2020s. *TechTrends, 67*(1), 12–25.
Kirschner, P. A., & Erkens, G. (2013). Toward a framework for CSCL research. *Educational Psychologist*, *48*(1), 1–8.
Koehler, M. J., Mishra, P., Kereluik, K., Shin, T. S., & Graham, C. R. (2013). The technological pedagogical content knowledge framework. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 101–111). New York: Springer.
Kovanovi , V., Gaševi , D., & Hatala, M. (2023). Generative AI and the future of learning analytics: Metacognition in the loop. *arXiv preprint*, arXiv:2307.02174.
Krueger, R. A., & Casey, M. A. (2014). *Focus groups: A practical guide for applied research* (5th ed.). Thousand Oaks, CA: SAGE.
Laurillard, D. (2012). *Teaching as a design science: Building pedagogical patterns for learning and technology.* London: Routledge.
Levin, I., Marom, M., & Kojukhov, A. (2025). Rethinking AI in Education: Highlighting the Metacognitive Challenge. *BRAIN: Broad Research in Artificial Intelligence and Neuroscience*, *16*(1 Suppl.1).
Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry.* Newbury Park, CA: SAGE.
Livingstone, S., & Sefton-Green, J. (2016). *The class: Living and learning in the digital age.* New York: NYU Press.
Luckin, R., & Holmes, W. (2016). *Intelligence unleashed: An argument for AI in education*. London: Pearson.
Mahamuni, A. J., Parminder, & Tonpe, S. S. (2024). Enhancing educational assessment with artificial intelligence: Challenges and opportunities**.** *IEEE Access, 12,* 1–10.
Manca, S., & Ranieri, M. (2017). Implications of social network sites for teaching and learning: Where we are and where we want to go. *Education and Information Technologies*, *22*, 605–622.
Mayring, P. (2014). *Qualitative content analysis: Theoretical foundation, basic procedures and software solution*. Klagenfurt: University of Klagenfurt.
McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica,* 22(3), 276–282.
Mercer, N. (2013). The social brain, language, and goal-directed collective thinking: A social conception of cognition and its implications for understanding how we think, teach and learn. *Educational Psychologist*, *48*(3), 148-168.
Mezirow, J. (1991). *Transformative dimensions of adult learning*. San Francisco, CA: Jossey-Bass.
Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Thousand Oaks, CA: SAGE Publications.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, *108*(6), 1017–1054.
MIUR. (2017). *Linee guida per l'inclusione scolastica degli alunni con disabilità*. Roma: Ministero dell'Istruzione, dell'Università e della Ricerca.
Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, *16,* 1-13.
OECD. (2018). *The future of education and skills: Education 2030*. Paris: OECD Publishing.
OECD. (2021). *AI and the future of skills*. Vol. 1. Paris: OECD Publishing.
Onwuegbuzie, A. J., & Leech, N. L. (2006). Linking research questions to mixed methods data analysis procedures. *The Qualitative Report*, *11*(3), 474–498.
Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). Thousand Oaks, CA:SAGE.
Plano Clark, V. L., & Ivankova, N. V. (2016). *Mixed methods research: A guide to the field*. Thousand Oaks, CA: SAGE.
Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, *12,* Article 22.
Prensky, M. (2001). Digital Natives, Digital Immigrants. *On the Horizon, 9*(5).
Rachbauer, T., Graup, J., & Rutter, E. (2025). Digital literacy and artificial intelligence literacy in teacher training. *Forum for Education Studies*, *3*(1), Article 1842.
Rashid, S. F., Duong-Trung, N., & Pinkwart, N. (2024). Generative AI in Education: Technical Foundations, Applications, and Challenges. In *Artificial Intelligence for Quality Education.* London: IntechOpen
Reich, J., & Ito, M. (2017). *From good intentions to real outcomes: Equity by design in learning technologies*. Irvine, CA: Digital Media and Learning Research Hub.
Røkenes, F. M., & Krumsvik, R. J. (2016). Prepared to teach ESL with ICT? A study of digital competence in Norwegian teacher education. *Computers & Education*, *97,* 1–20.
Rubach, C., & Lazarides, R. (2021). Addressing 21st-century digital skills in schools—Development and validation of an instrument to measure teachers' basic ICT competence beliefs. *Computers in Human Behavior, 118*, Article 106636.
Sadykova, G., & Kayumova, A. (2024). Educators' perception of artificial intelligence as instructional tool. *TEM Journal*, *13*(4), 3194–3204.
Saldaña, J. (2021). *The coding manual for qualitative researchers* (4th ed.). London: SAGE.
Sánchez, E. E., Aguilar Romero, P. G., Hidalgo Durán, A. E., Ramírez Vargas, A. del R., Hidalgo Durán, J. A., & Guamán Chimbo, E. E. (2025). Artificial intelligence for inclusive classrooms: Transforming learning through innovation and digital equity. *Revista Multidisciplinar de Estudios*, *4*(4), 595–615.
Scherer, R., Siddiq, F., & Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education, 128*, 13–35.
Schleicher, A. (2018). *World Class: How to Build a 21st-Century School System.* Paris: OECD Publishing.
Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.
Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Cambridge: Polity Press.
Sperling, K., Stenberg, C.-J., McGrath, C., Åkerfeldt, A., Heintz, F., & Stenliden, L. (2024). In search of artificial intelligence (AI) literacy in teacher education: A scoping review. *Computers and Education Open, 6*, 100169.
Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks, CA: SAGE.
Tondeur, J., Scherer, R., Baran, E., & Sointu, E. (2019*).* Teacher educators as gatekeepers: Preparing the next generation of teachers for technology integration in education*. British Journal of Educational Technology*, *50*(3), 1189–1209.
Tondeur, J., van Braak, J., Sang, G., Voogt, J., Fisser, P., & Ottenbreit-Leftwich, A. (2012). Preparing pre-service teachers to integrate technology: A synthesis of qualitative evidence. *Computers & Education*, *59*(1), 134–144.
Tracy, S. J. (2010). Qualitative quality: Eight "big-tent" criteria. *Qualitative Inquiry, 16*(10), 837–851.
Tsirantonaki, S., & Vlachou, A. (2025). AI-driven inclusive practices: Innovative approaches to differentiated teaching and educational inclusion. *European Journal of Inclusive Education*, *4*(1), 134–146.
Vygotskij, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
Vuorikari, R., Kluzer, S., & Punie, Y. (2022). *DigComp 2.2: The digital competence framework for citizens – With new examples of knowledge, skills and attitudes*. Luxembourg: Publications Office of the European Union.
Whitelock, D., & Cross, S. (2012). Authentic assessment: What does it mean and how is it instantiated by a group of distance learning academics? *International Journal of e-Assessment*, *2*(1).

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, *45*(3), 223–235.

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.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on AI in education – where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 39.