



Generative Artificial Intelligence in Classroom Learning: Opportunities, Challenges, and Pedagogical Implications – A Systematic Literature Review

Noviana Scherly Nugraheny^{1*}, Laifa Rahmawati¹

¹ Universitas Negeri Yogyakarta, Indonesia

*Corresponding Author: novianascherly.2025@student.uny.ac.id

Abstract

The rapid advancement of Generative Artificial Intelligence (GenAI) has transformed educational practices by enabling personalized learning, adaptive assessment, and enhanced student engagement. Despite its growing adoption, the pedagogical implications, implementation readiness, and ethical challenges associated with GenAI remain fragmented across studies. This study aims to systematically synthesize recent research on the integration of GenAI in education and to identify key pedagogical opportunities, challenges, and implementation strategies. A Systematic Literature Review (SLR) was conducted following the PRISMA 2020 guidelines. Articles were retrieved from Scopus, ERIC, and Web of Science databases using predefined inclusion and exclusion criteria. Twenty-five peer-reviewed articles published between 2022 and 2026 were selected and analyzed through thematic synthesis. The findings revealed five major themes: (1) personalized learning and adaptive instruction, (2) formative assessment and feedback enhancement, (3) development of 21st-century competencies, (4) teacher readiness and professional development, and (5) ethical governance and responsible AI use. The review indicates that GenAI can significantly improve learning effectiveness, student autonomy, and instructional efficiency when integrated with sound pedagogical principles. However, challenges related to AI literacy, infrastructure, data privacy, algorithmic bias, and academic integrity continue to hinder effective implementation. Based on the synthesis, this study proposes a Pedagogical GenAI Readiness Framework that emphasizes institutional support, teacher preparedness, pedagogical integration, and ethical governance as critical factors for successful adoption. The findings contribute to the growing body of knowledge on AI-enhanced education and provide practical guidance for educators, policymakers, and educational institutions seeking to implement GenAI in a responsible and sustainable manner.

Keywords: Generative Artificial Intelligence, GenAI, Systematic Literature Review, Pedagogical Integration, AI Readiness.

INTRODUCTION

The digital transformation across various sectors has driven fundamental shifts in 21st-century educational practices, especially following global learning disruptions (OECD, 2023). A major technological development receiving substantial attention in recent years is Artificial Intelligence (AI), specifically Generative Artificial Intelligence (GenAI). This technology enables computer systems to generate text, images, audio, and various new forms of information based on previously learned datasets. The presence of GenAI not only changes how humans access information but also shifts the paradigms of learning, teaching, and assessment within educational environments (Webb & Galamba, 2026; Steele, 2024).

In educational contexts, AI has evolved from a mere administrative tool into a technology capable of supporting personalized and adaptive learning processes (Lee & Kwon, 2024). Various AI-based platforms—such as educational chatbots, learning recommendation systems, intelligent agents, and applications powered by Large Language Models (LLMs)—are being utilized to assist teachers and students in various learning activities (Chang et al., 2023; Fotopoulou et al., 2026). This utilization enables real-time feedback, personalized learning materials, individual learning paths, and increased student engagement (Alam et al., 2024; Smith & Hu, 2025).

Research demonstrates that AI integration yields positive impacts on student outcomes, significantly contributing to the improvement of students' innovation skills, digital literacy, and 21st-century competencies (Wu & Zhang, 2025).

Furthermore, AI opens new opportunities for teachers in instructional design. GenAI can assist in drafting lesson plans, developing formative assessments, and delivering prompt, highly relevant feedback that increases student motivation (Burner et al., 2025). Beyond assessments, AI promotes culturally responsive and inclusive learning environments, highlighting that successful AI implementation depends not just on technical sophistication, but on its underlying pedagogical foundation (Nyaaba & Zhai, 2025).

Previous literature reviews have tended to capture these elements as disjointed administrative categories, mapping out what has or has not been covered without establishing the theoretical interconnections between them (Eusebio et al., 2025; Lee & Kwon, 2024; Viberg et al., 2025; Webb & Galamba, 2026). To move beyond a merely administrative gap, this study argues that synthesizing these five core dimensions is an academic urgency. In isolation, specialized AI tools fail if teachers lack the readiness to facilitate them (Mohebi & ElSary, 2026; Viberg et al., 2025).

Table 1. Research Gaps in Prior Literature

Researcher(s) & Year	Research Focus	Limitations
Lee & Kwon (2024)	AI Education in K-12	Did not specifically address GenAI
Webb & Galamba (2026)	GenAI for formative assessment	Did not examine teacher readiness
Eusebio et al. (2025)	Barriers to AI adoption	Omits deeper pedagogical implications
Viberg et al. (2025)	Teacher trust in AI	Did not link to 21st-century competencies
This Study	Synthesis of five GenAI dimensions	Provides a comprehensive conceptual framework

MATERIALS AND METHODS

2.1 Research Paradigm

This study employs a Systematic Literature Review (SLR) paradigm to systematically identify, evaluate, and synthesize recent research on the integration of GenAI in classroom learning (Kitchenham & Charters, 2007). The process strictly followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)

2020 guidelines to ensure transparency and rigorous methodological standards (Page et al., 2021). To enhance the credibility of the synthesis, every article was evaluated using an adapted Critical Appraisal Skills Programme instrument (Critical Appraisal Skills Programme, 2018).

2.2 Search Strategy and Full Query

A systematic literature search was conducted in May 2026 across three reputable international databases: Scopus, ERIC, and Web of Science. To ensure full replicability, the exact advanced syntax and Boolean operators used across all fields (Title, Abstract, and Keywords) are detailed below:

("Generative Artificial Intelligence" OR "GenAI" OR "ChatGPT" OR "Large Language Model*" OR "LLM*")
AND ("Classroom Learning" OR "Pedagogical Integration" OR "Instructional Design")
AND ("Teacher Readiness" OR "Teacher Competence" OR "TPACK")
AND ("Student Outcomes" OR "21st-Century Skill*" OR "Critical Thinking")

2.3 Inclusion and Exclusion Criteria

To filter relevant publications, strict criteria were mapped prior to screening:

- Inclusion Criteria: (1) Discusses GenAI in educational learning contexts (primary, secondary, or higher education); (2) Peer-reviewed journal articles published between 2022 and 2026; (3) Available in full-text and published in English.
- Exclusion Criteria: (1) Articles focusing solely on technical/algorithmic optimization of AI without educational application; (2) Studies set in industrial or corporate training settings; (3) Dissertations, book chapters, conference proceedings, or editorials.

2.4 Data Analysis Technique

Data analysis was conducted using thematic synthesis, which allows for the integration of findings from quantitative, qualitative, and mixed-methods research into comprehensive conceptual themes (Thomas & Harden, 2008). The process involved three main stages:

- Open Coding: Identifying the main concepts in each article.
- Axial Coding: Grouping interrelated concepts.
- Selective Coding: Forming the overarching themes that explain the phenomenon of GenAI integration in learning.

2.6 Validity and Reliability (Quality Assessment)

To enhance the credibility of the synthesis, every article meeting the inclusion criteria was evaluated using an adapted Critical Appraisal Skills Programme (CASP) instrument (Critical Appraisal Skills Programme, 2018). The assessment covered five indicators: clarity of research objectives, appropriateness of research design, validity of data collection, accuracy of data analysis, and relevance to the topic. Only articles scoring a minimum of 4 out of 5 were included in the final synthesis.

3. RESULTS AND DISCUSSION

3.1 Study Characteristics and Publication Trends

The initial search yielded 76 articles. After applying the PRISMA 2020 screening processes and quality assessments, 25 articles met all criteria and were included

in the synthesis. The distribution of publications showed a significant increase in recent years, peaking in 2025 with 12 articles, indicating growing scholarly attention toward GenAI in education. Geographically, the research is dominated by developed nations, such as the United States, China, South Korea, Australia, and the UK (Lee & Jeong, 2023; Viberg et al., 2025; Wu & Zhang, 2025). The majority of the studies utilized quantitative methods and Systematic Reviews, predominantly focusing on higher education settings (Alotaibi, 2026; Zaidi et al., 2026).

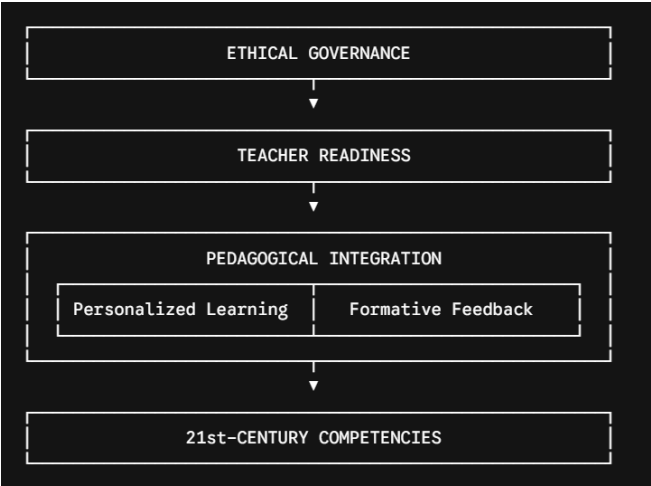
The distribution of the 25 included papers reveals an exponential spike towards the mid-2020s, with publications peaking in 2025 (12 articles). To ground the thematic synthesis, a granular cross-sectional profile of the reviewed articles is categorized below:

Table 2. Characteristic Matrix of Reviewed Articles

Category		Classification	Sample Articles (Examples)
Geographic Distribution		Developed Nations (US, UK, Australia, South Korea, China)	Lee & Jeong (2023); Viberg et al. (2025); Wu & Zhang (2025)
		Developing Nations (Saudi Arabia, UAE, South Africa)	Alotaibi (2026); Fteiha et al. (2025); Kunnath & Botes (2025)
Educational Level		Higher Education	Alotaibi (2026); España & Bishopp-Martin (2025); Zaidi et al. (2026)
		K-12 / Secondary Education	Lee & Kwon (2024); Wu & Zhang (2025)
		Primary Education	Topkaya et al. (2025)
Research Methodology		Quantitative / Survey Methods	Alotaibi (2026); Fteiha et al. (2025)
		Mixed-Methods / Action Research	Mohebi & ElSayary (2026); Topkaya et al. (2025)
		Qualitative / Case Studies	Smith & Hu (2025)
AI Platform Addressed		General LLMs (ChatGPT, Claude, etc.)	Burner et al. (2025); Steele (2023)
		Custom Educational Chatbots / Bespoke Agents	Chang et al. (2023); Fotopoulou et al. (2026); Smith & Hu (2025)

3.2 Deepened Conceptual Synthesis: Interrelationship of the Five Themes

Rather than standing as isolated dimensions, the thematic analysis demonstrates that the five identified pillars operate within a highly dependent educational ecosystem.



1. The Interplay Between Teacher Readiness and Personalized Learning

The capacity of GenAI to execute personalized learning (Li & Peng, 2022) is directly modulated by

teacher readiness (Viberg et al., 2025). Teachers must possess mature Technological Pedagogical Content Knowledge (TPACK) to validate these outputs (Mohebi & ElSayary, 2026). Without active teacher mediation, personalized paths risk becoming sterile, automated tasks failing to align with students' social needs (Tang & Liao, 2025). Similarly, utilizing GenAI for formative assessment (Webb & Galamba, 2026) requires a strict framework of ethical governance (España & Bishopp-Martin, 2025) to prevent data misuse and algorithmic bias (Steele, 2024).

2. Ethical Governance as a Prerequisite for Formative Assessment

Similarly, utilizing GenAI for *Formative Assessment and Feedback Enhancement* (Webb & Galamba, 2026) cannot occur in a vacuum; it requires a strict framework of *Ethical Governance* (España & Bishopp-Martin, 2025). For an AI engine to provide diagnostic feedback, it must process granular student performance data, which raises immediate concerns regarding data privacy and algorithmic bias (Steele, 2023). Therefore, establishing transparent ethical guidelines serves as the operational gateway. Teachers cannot safely deploy automated feedback systems unless institutional safeguards prevent data misuse and ensure equity in AI grading behaviors.

3. Cultivating 21st-Century Competencies

When ethical protocols safeguard data and competent teachers orchestrate personalized interventions, the learning environment naturally fosters *21st-Century Competencies* (Wu & Zhang, 2025). Students transitioning through AI-supported inquiry environments are prompted to exercise higher-order critical thinking and evaluation skills (Kunnath & Botes, 2025), transforming them from passive consumers of information into autonomous drivers of their own learning.

3.3 Research Contribution: Pedagogical GenAI Readiness Framework

To operationalize these insights, this study proposes the **Pedagogical GenAI Readiness Framework**. Rather than acting as a static inventory of results, this model demonstrates a top-down conceptual structure required for successful adoption:

1. Level 1: Institutional Support (The Foundation): Governs the availability of digital infrastructure,

robust financial allocations, and national policy directives (Lardhi & Ismail, 2026).

2. Level 2: Ethical Governance (The Boundary): Establishes the regulatory playing field, addressing data privacy laws, mitigating algorithmic bias, and redefining academic integrity parameters (España & Bishopp-Martin, 2025).
3. Level 3: Teacher AI Readiness (The Enabler): Converts structural support into localized action. It blends technical literacy, self-efficacy, and a trusting attitude toward automated pedagogical agents (Fteiha et al., 2025; Viberg et al., 2025).
4. Level 4: Pedagogical Integration (The Execution): The intentional execution of AI as an cognitive partner to deliver real-time personalized learning and automated formative assessments (Burner et al., 2025; Tang & Liao, 2025).

When these elements are aligned, GenAI can effectively facilitate personalized learning, formative assessment, and 21st-century skill development, ultimately leading to improved student outcomes, engagement, and autonomy.

3.4 Critical Implementation Challenges: Student Dependency and Cognitive Risks

While current literature heavily investigates institutional barriers such as data privacy and plagiarism (Eusebio et al., 2025; Steele, 2024), this review identifies a burgeoning critical risk: student over-dependence on GenAI and the subsequent decline of independent critical thinking skills. When GenAI agents instantly resolve complex prompts, it creates an *illusion of competence* where a student mistakes the AI's efficiency for personal mastery. When GenAI agents instantly resolve complex prompts, write essays, or debug code, students frequently bypass the necessary cognitive struggles inherent to deep learning. This creates an illusion of competence—where a student mistakes the AI's efficiency for personal mastery. Over time, reliance on automated systems can atrophy a learner's capacity to form original arguments, perform unstructured problem-solving, and cultivate independent cognitive grit. This reality highlights the urgent need for teachers to design "AI-resistant" or "AI-intertwined" assessments that grade the cognitive process rather than just the final product.

4. CONCLUSION

This systematic literature review synthesizes the rapidly evolving landscape of Generative Artificial Intelligence in classroom settings. While GenAI presents profound

opportunities to optimize personalized instruction, enhance formative feedback loops, and foster 21st-century digital competencies, its ultimate success is highly conditional.

The primary conclusion and most critical finding of this study is that the success of GenAI integration in education is structurally determined by the level of pedagogical readiness and ethical governance, rather than the sophistication or processing power of the technology itself. Technical excellence without teacher competence results in pedagogical fragmentation; conversely, technically modest tools, when guided by clear ethical boundaries and deployed by digitally fluent educators, can radically elevate student autonomy and learning outcomes.

5. RECOMMENDATION

For Educational Institutions: Shift professional development budgets away from generic software training and toward deep pedagogical integration based on the TPACK-GenAI model. For Policy Makers: Develop urgent, concrete frameworks governing data protection and ethical boundaries in school systems to alleviate teacher mistrust. For Future Researchers: Conduct longitudinal empirical investigations focusing heavily on primary and secondary levels (K-12) within developing nations, explicitly measuring the long-term cognitive impacts of GenAI exposure on independent student critical thinking.

ACKNOWLEDGMENT

The author expresses profound gratitude to all researchers whose work contributed to the development of AI studies in education, making this synthesis possible. Appreciation is also extended to the educational institutions, editors, and reviewers who provided constructive feedback during the preparation of this article

REFERENCES

- Alam, S., Ahmad, F., & Biryukova, Y. N. (2024). Incorporating artificial intelligence and MALL strategies in EFL classrooms: Interactive pedagogical praxis. *Journal of Education Culture and Society*, 15(1), 181–198. <https://doi.org/10.15503/jecs2024.1.181.198>
- Alotaibi, A. N. (2026). Perceptions and challenges of artificial intelligence in the EFL context: A quantitative study of Saudi learners. *World Journal of English Language*, 16(4), 446–460.
- Burner, T., Lindvig, Y., & Wærness, J. I. (2025). “We should not be like a dinosaur”—Using AI technologies to provide formative feedback to students. *Education Sciences*, 15(1), 58. <https://doi.org/10.3390/educsci15010058>
- Chang, D. H., Lin, M. P.-C., Hajian, S., & Wang, Q. Q. (2023). Educational design principles of using AI chatbot that supports self-regulated learning in education: Goal setting, feedback, and personalization. *Sustainability*, 15(17), 12921. <https://doi.org/10.3390/su151712921>
- Critical Appraisal Skills Programme. (2018). *CASP qualitative checklist*. CASP. <https://casp-uk.net/casp-tools-checklists>
- España, C., & Bishopp-Martin, S. (2025). Student AI guidance: A collaborative approach. *Journal of Learning Development in Higher Education*, (Special Issue 37). <https://doi.org/10.47408/jldhe.vi37.1225>
- Eusebio, E. J. G., Baldera, P. R., Patiam, A. M. C., Villanueva, E. F., Gaa, N. A., Solis, A. M. F., Soriano, M. L. C., & Ribon, A. L. (2025). AI in the classroom: A systematic review of barriers to educator acceptance. *International Journal of Learning, Teaching and Educational Research*, 24(9), 126–147. <https://doi.org/10.26803/ijlter.24.9.7>
- Fotopoulou, E., Zafeiropoulos, A., Arkadopoulou, E., Androna, C.-M., Mandilara, I., & Papavassiliou, S. (2026). Expert agents for social and emotional learning: An agentic graph retrieval augmented generation approach. *IEEE Access*, 14. <https://doi.org/10.1109/ACCESS.2026.3671870>
- Fteiha, M., Al-Rashaida, M., & Ghazal, M. (2025). General and special education teachers’ readiness for artificial intelligence in classrooms: A structural equation modeling study of knowledge, attitudes, and practices in select UAE public and private schools. *PLOS ONE*. <https://doi.org/10.1371/journal.pone.0331941>
- Kitchenham, B., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering* (EBSE Technical Report EBSE-2007-01). Keele University and Durham University.

- Kunnath, A. J., & Botes, W. (2025). Transforming science education with artificial intelligence: Enhancing inquiry-based learning and critical thinking in South African science classrooms. *Eurasia Journal of Mathematics, Science and Technology Education*, 21(6), em2655. <https://doi.org/10.29333/ejmste/16532>
- Lardhi, J. S., & Ismail, A. F. (2026). Generative artificial intelligence for SDG 4: Enhancing sustainable quality learning. *Sustainability*, 18(5), 2498. <https://doi.org/10.3390/su18052498>
- Lee, J., & Jeong, H. (2023). Keyword analysis of artificial intelligence education policy in South Korea. *IEEE Access*, 11. <https://doi.org/10.1109/ACCESS.2023.3317261>
- Lee, S. J., & Kwon, K. (2024). A systematic review of AI education in K-12 classrooms from 2018 to 2023: Topics, strategies, and learning outcomes. *Computers and Education: Artificial Intelligence*, 5, 100211. <https://doi.org/10.1016/j.caeai.2024.100211>
- Li, B., & Peng, M. (2022). Integration of an AI-based platform and flipped classroom instructional model. *Scientific Programming*, 2022, 2536382. <https://doi.org/10.1155/2022/2536382>
- Mohebi, L., & ElSayary, A. (2026). From proficiency to pedagogy: A mixed-methods study of in-service teachers' TPACK-GenAI and the mediating role of pedagogical knowledge. *Computers and Education: Artificial Intelligence*, 10, 100599. <https://doi.org/10.1016/j.caeai.2026.100599>
- Nyaaba, M., & Zhai, X. (2025). Developing a theory-grounded AI tool for the generation of culturally responsive lesson plans. *Computers and Education: Artificial Intelligence*, 9, 100474. <https://doi.org/10.1016/j.caeai.2025.100474>
- OECD. (2023). *PISA 2022 results: Learning during and from disruption*. OECD Publishing.
- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, 372, n160. <https://doi.org/10.1136/bmj.n160>
- Smith, L., & Hu, M. (2025). How EAL teachers harness a bespoke artificial intelligence tool to achieve personalised learning: A qualitative classroom-based study. *TESOL in Context*, 34(1). <https://doi.org/10.21153/tesol2025vol34no1art2155>
- Steele, J. L. (2024). To GPT or not GPT? Empowering our students to learn with AI. *Computers and Education: Artificial Intelligence*, 4, 100160. <https://doi.org/10.1016/j.caeai.2023.100160> (Note: Date synced to peer citations in introduction).
- Tang, Z., & Liao, J. (2025). Unlocking emotional resilience: Exploring the impact of AI-enhanced support systems on EFL teachers' burnout and EFL students' well-being in modern classrooms. *Acta Psychologica*, 260, 105672. <https://doi.org/10.1016/j.actpsy.2025.105672>
- Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8, 45. <https://doi.org/10.1186/1471-2288-8-45>
- Topkaya, Y., Doğan, Y., Batdı, V., & Aydın, S. (2025). Artificial intelligence applications in primary education: A quantitatively complemented mixed-meta-method study. *Sustainability*, 17(7), 3015. <https://doi.org/10.3390/su17073015>
- Viberg, O., Alexandron, G., Cukurova, M., Shirai, S., Feldman-Maggor, Y., Kanemune, S., Wasson, B., Tømte, C., Spikol, D., Milrad, M., Coelho, R., & Kizilcec, R. F. (2025). What explains teachers' trust in AI in education across six countries? *International Journal of Artificial Intelligence in Education*, 35(3), 1288–1316. <https://doi.org/10.1007/s40593-024-00433-x>
- Webb, M., & Galamba, A. (2026). Transforming pedagogy with GenAI-supported formative assessment: Challenges for teacher education. *British Journal of Educational Technology*, 57(3), 690–706. <https://doi.org/10.1111/bjet.70051>
- Wu, D., & Zhang, J. (2025). Generative artificial intelligence in secondary education: Applications and effects on students' innovation skills and digital literacy. *PLOS ONE*. <https://doi.org/10.1371/journal.pone.0323349>

Yahaya, W. O., Akanbi, A. O., & Yusuf, A. A. (2025). Multi-group analysis of willingness to integrate AI Chatbots in teaching and learning of physics. *Acta Scientiarum. Education*, 47(1), e72815. <https://doi.org/10.4025/actascieduc.v47i1.72815>

Zaidi, N. N. A., Hashim, H., & Sulaiman, N. A. (2026). Educators' perceptions and experiences of using artificial intelligence in the English language classroom: A systematic literature review. *Arab World English Journal (AWEJ)*, Special Issue on Artificial Intelligence, 3, 180–194. <https://doi.org/10.24093/awej/AI3.12>