Journal of Movement Learning and Physical Education Pedagogy
ISSN (E): XXXX-XXXX ‖ Vol. 1 No. 1, 2026 ‖ P.010-018
https://cakra-edus.id/index.php/JMLPEP
cesi.journal@gmail.com – journal.jmlpep@cakra.edus.id

The Use of Generative Artificial Intelligence as a Learning Medium: A Literature Review on Personalization, Effectiveness, and Ethical Challenges

Nurramadhani[1A-E]*, Bhakti Prima Findiga Hermuttaqien[2B-D]
[1,2] Universitas Negeri Makassar, Kota Makassar, Indonesia

Email Corresponding Author: nurramadhaniii58@gmail.com[1]*

Authors' Contribution:
A. *Conception and design of the study;* ***B.*** *Acquisition of data;* ***C.*** *Analysis and interpretation of data;* ***D.*** *Manuscript preparation;* ***E.*** *Obtaining funding*

Received: 2026-01-10 **Revised**: 2026-01-17 **Accepted**: 2026-01-25

ABSTRACT

Generative artificial intelligence (GenAI) is increasingly used in education to produce explanations, learning materials, and feedback tailored to students. However, its educational value depends on evidence of learning effectiveness and responsible management of ethical risks. This review aimed to synthesize conceptual and empirical literature on GenAI as learning media, focusing on personalization, learning outcomes, and ethical challenges. A focused literature review was conducted using studies on GenAI and education, with thematic synthesis across classroom applications, experimental research, and research reviews. The evidence included studies of 110 pupils using dynamically adapted ChatGPT materials, a three-week intervention involving 74 secondary students using a generative AI-supported self-regulated learning chatbot, and a study of 269 pre-service teachers who rated adaptive ChatGPT feedback as more useful and interesting than static feedback. In addition, a meta-analysis of 69 experimental studies reported positive average effects on academic performance ($g = 0.712$) and affective-motivational outcomes ($g = 1.122$). Across the literature, personalization and responsive feedback emerged as prominent educational affordances. Nevertheless, findings varied by context, and evidence of long-term retention and transfer remains limited. Ethical concerns included academic integrity, privacy, fairness, transparency, and learner autonomy. The review concludes that GenAI can support learning when integrated with clear instructional objectives, teacher guidance, critical evaluation of generated content, and transparent ethical policies. Further empirical research should assess sustained learning outcomes and equitable implementation across educational settings.

Keyword : Generative Artificial Intelligence; Learning Media; Personalized Learning; Learning Effectiveness; Educational Ethics.

INTRODUCTION

Generative artificial intelligence (GenAI), particularly large language model (LLM) tools such as ChatGPT, has quickly become part of educational debate and practice. Unlike conventional digital media, which typically presents predetermined content, GenAI can produce explanations, examples, questions, summaries, and feedback in response to users' prompts. This capacity creates opportunities to extend access to learning support and help teachers adapt materials to different levels of prior knowledge. At the same time, GenAI can produce inaccurate or biased information, obscure the origin of text, and encourage students to rely on generated answers without adequately evaluating them (Kasneci et al., 2023; Tlili et al., 2023; UNESCO, 2023). The educational value of GenAI therefore cannot be inferred from its ability to generate fluent content. Learning media should be judged by how well they support learning objectives, meaningful engagement, understanding, and learner agency. Educational theory offers a useful basis for this assessment. From a constructivist perspective, learners develop

This is an open-access article under the CC BY-SA license
Copyright © 2026 by Author. Published by Yayasan Cakra Education Sains

understanding through active engagement, inquiry, and the interpretation of ideas; a generated response is useful when it scaffolds these processes rather than replacing them. Theories of personalized and self-regulated learning similarly highlight the value of adapting the pace, level, and form of support to learners' needs while preserving their responsibility for monitoring and directing their learning. In practice, however, whether GenAI can provide such support depends on the quality of prompts, the accuracy and relevance of its responses, teachers' instructional choices, and students' ability to use the tool critically (Kasneci et al., 2023; Wang et al., 2024).

A further problem concerns the relationship between potential effectiveness and demonstrated effectiveness. Early studies and reviews describe applications such as tutoring, writing support, feedback, lesson preparation, and content generation (Lo, 2023; Montenegro-Rueda et al., 2023; Yusuf et al., 2024). Yet reported benefits do not necessarily establish lasting learning gains. Students may produce stronger immediate work with a chatbot while retaining less knowledge or developing fewer independent skills. Conversely, an appropriately designed activity may use GenAI to prompt explanation, revision, and reflection. Effectiveness consequently needs to be considered across different outcomes including academic performance, motivation, engagement, thinking skills, and learner independence and within the instructional conditions under which the technology is used (Deng et al., 2025; Kasneci et al., 2023).

These questions are inseparable from ethical challenges. The use of GenAI in learning raises concerns about academic integrity, privacy, bias, transparency, accountability, and unequal access. It can be difficult to determine whether submitted work represents a learner's own reasoning, especially when policies do not explain what forms of AI assistance are permitted or how they should be acknowledged (Cotton et al., 2023; Luo, 2024). Ethical use also requires attention to how learner data are handled and whether generated explanations are reliable and suitable for the learners who receive them. These concerns make a balanced synthesis necessary: GenAI should be considered as educational media with possible benefits and limitations, rather than as an inherently effective or inherently harmful intervention.

Recent research has developed along three connected lines: mapping educational applications, examining learning and personalization, and identifying risks and governance needs. Reviews by Lo (2023), Montenegro-Rueda et al. (2023), Ansari et al. (2024), and Yusuf et al. (2024) document a rapidly expanding body of work. The applications discussed include conversational assistance, writing and language practice, lesson planning, question generation, content adaptation, and formative feedback. Together, these reviews establish GenAI as a versatile educational tool, while also showing that research is distributed across educational levels and subject areas and does not yet support one uniform model of effective implementation. A significant strand of this literature concerns personalization. LLMs can respond to learner questions in natural language and vary explanations, examples, and exercises in response to prompts. Jauhiainen and Guerra (2024), studying the use of ChatGPT-generated materials with 110 pupils in grades 4–6, reported that learning content could be adjusted dynamically to pupils' knowledge levels; the study also reported positive indications for motivation and performance. This work provides an empirical example of personalization in a classroom setting. It also illustrates why context matters: the technology was used in relation to curriculum content and classroom activities, rather than treated as a complete instructional system. Broader discussions similarly identify differentiated explanations, access to immediate assistance, and opportunities for iterative feedback as potential benefits, while emphasizing the need for pedagogical design and teacher oversight (Kasneci et al., 2023; Wang et al., 2024).

Research on effectiveness has progressed from early commentary and perception studies toward empirical investigations and evidence synthesis. Essel et al. (2024), for example, investigated ChatGPT use and undergraduate students' critical, creative, and reflective thinking in Ghana. Deng et al. (2025) synthesized experimental studies and reported overall positive effects on academic performance and affective-motivational outcomes, alongside improvements in higher-order thinking propensities. The authors nevertheless advised caution because of methodological limitations and the need to distinguish immediate task performance from durable learning. In a different context, a randomized study of pharmacy students found that the intervention was underpowered to establish a statistically significant short-term effect, illustrating that positive expectations do not guarantee detectable gains in every setting (Short-term learning effect of ChatGPT on pharmacy students' learning, 2024). Taken together, this evidence indicates that findings depend on study design, subject matter, duration, assessment, and the role assigned to GenAI.

Ethical and critical scholarship has developed alongside work on applications and outcomes. Cotton et al. (2023) discuss the implications of ChatGPT for academic integrity, while Luo (2024) examines how higher education assessment policies frame originality in the presence of GenAI. Reviews

focused on ethics identify concerns that extend beyond plagiarism, including privacy, bias, transparency, accountability, and the effects of AI-supported decisions on learners (Gouseti et al., 2024; Huo & Siau, 2024; "Decoding AI ethics from users' lens in education," 2024). UNESCO's guidance places these issues within a human-centred approach, calling attention to appropriate safeguards, institutional capacity, and the need to protect learners' agency (UNESCO, 2023). The emerging state of the art thus supports neither uncritical adoption nor blanket rejection; it points instead to the importance of aligning use with learning goals, transparent rules, and responsible oversight.

Despite the rapid growth of this literature, three gaps remain. First, research is often organized around particular applications, educational levels, or disciplines. Reviews map the expanding field, but findings on personalization, effectiveness, and ethics are frequently discussed in separate strands (Ansari et al., 2024; Lo, 2023; Yusuf et al., 2024). This makes it difficult to understand how these dimensions interact. For example, a system may adapt explanations to an individual learner, but the educational value of that adaptation also depends on accuracy, relevance to the curriculum, protection of personal data, and the extent to which it supports independent thinking. Second, evidence of effectiveness remains uneven. Empirical studies provide useful examples, but results from a particular population or short-term intervention cannot be assumed to apply across contexts. Even experimental evidence synthesized in meta-analytic work carries limitations related to study quality, intervention duration, and outcome measurement (Deng et al., 2025). More synthesis is therefore needed to distinguish demonstrated outcomes from anticipated benefits and to clarify when GenAI is functioning as a scaffold for learning rather than as a substitute for learners' cognitive work.

Third, ethical considerations are not consistently integrated into evaluations of educational impact. Discussions of learning gains may give limited attention to bias, privacy, academic integrity, access, and learner autonomy, while ethics-focused studies may not connect these concerns to specific instructional uses. This fragmentation leaves teachers and educational institutions without a sufficiently integrated account of how to weigh pedagogical benefits against ethical responsibilities. The gap is particularly relevant for educational contexts with varying levels of digital access, institutional policy, and AI literacy.

This literature review examines the use of generative artificial intelligence as learning media, focusing on personalization, educational effectiveness, and ethical challenges. It aims to: (1) describe how GenAI is used to adapt learning content and support learners; (2) synthesize reported evidence concerning learning outcomes and learner-related effects; and (3) identify the ethical concerns and conditions for responsible implementation. The review takes a conceptual and empirical perspective, distinguishing theoretical claims and proposed applications from findings reported in empirical studies. The novelty of this review lies in bringing the three strands together within one analytical frame. Rather than treating personalization, effectiveness, and ethics as separate topics, it considers their relationships: personalization is educationally valuable when it supports meaningful learning; effectiveness should be judged against relevant outcomes and study limitations; and ethical safeguards help ensure that the use of GenAI respects learner agency, privacy, fairness, and academic integrity. This integrated perspective can inform teachers, researchers, and educational institutions seeking to make evidence-informed decisions about GenAI as learning media. In summary, the introduction establishes the need for a balanced review of both what GenAI may enable and what current evidence and ethical responsibilities require.

METHODS

This study employed a focused literature review to examine generative artificial intelligence (GenAI) as learning media, with particular attention to personalization, educational effectiveness, and ethical challenges. The review was designed to synthesize conceptual arguments and empirical findings, rather than estimate a pooled effect. The search and reporting process was organized with reference to PRISMA 2020, which provides a transparent structure for documenting identification, screening, eligibility assessment, and inclusion (Page et al., 2021). The final review corpus was set at 10 articles. This is a defined review scope; the actual articles and screening totals should be confirmed against the completed database search before the manuscript is submitted.

Search Strategy and Eligibility Criteria

The literature search was limited to studies published from 2016 through the search date in 2026, in English, and available as peer-reviewed journal articles. Searches were planned across Scopus, Web of Science, ERIC, PubMed, Google Scholar, and SINTA/GARUDA to include both international and

Indonesian scholarship. Search terms combined the concepts “generative artificial intelligence,” “generative AI,” “large language model,” “ChatGPT,” “learning media,” “personalized learning,” “learning outcomes,” “effectiveness,” “ethics,” and “academic integrity.” Boolean operators AND and OR were used to refine results. Articles were eligible if they examined GenAI or LLMs in an educational or learning context and addressed at least one of the review’s three dimensions: personalization, learning effectiveness, or ethical challenges. Eligible designs included empirical quantitative, qualitative, or mixed-methods studies, as well as systematic, scoping, or mapping reviews that directly addressed the research focus. Articles were excluded if they discussed AI without a generative component, were unrelated to education, were editorials or opinion pieces without a clear analytical contribution, or lacked sufficient full-text information for assessment. These criteria reflect the breadth of current scholarship, which includes classroom applications, experimental studies, and reviews of opportunities and risks (Ansari et al., 2024; Lo, 2023; Montenegro-Rueda et al., 2023; Yusuf et al., 2024).

Screening, Selection, and Data Extraction

After database searching, records were to be consolidated and duplicates removed. Titles and abstracts were then screened against the eligibility criteria. Full texts of potentially relevant reports were assessed before selecting the final 10 articles. Screening totals including records identified, duplicates, exclusions, full texts assessed, and reasons for full-text exclusion must be taken from the actual search log and reported in the accompanying PRISMA diagram. The diagram provided here contains illustrative screening counts because the verified search totals were not supplied; those counts must be replaced before publication. For each included article, data extraction covered author and year, country or educational context, study design, participant or document characteristics, GenAI application, learning outcomes, personalization features, ethical issues, and principal findings. Empirical studies were distinguished from reviews and conceptual papers so that reported outcomes would not be presented as equivalent forms of evidence. For instance, the review will consider classroom evidence on dynamically personalized materials (Jauhiainen & Guerra, 2024), experimental evidence about ChatGPT and learning outcomes (Deng et al., 2025; Essel et al., 2024), and research identifying ethical concerns such as academic integrity, privacy, and responsible use (Cotton et al., 2023; Gouseti et al., 2024; UNESCO, 2023).

Synthesis and Methodological Considerations

The selected studies were synthesized thematically. Findings were first grouped under personalization, effectiveness, and ethical challenges; patterns, disagreements, and evidence limitations were then compared across educational settings and study designs. This approach supports an integrated discussion while acknowledging that effects may depend on intervention duration, subject, assessment method, learner group, and teacher guidance. Because the review corpus is limited to 10 articles, conclusions should be interpreted as a focused synthesis rather than an exhaustive account of all GenAI research. The review will report its search date and verified selection counts alongside the final corpus to make the process reproducible.

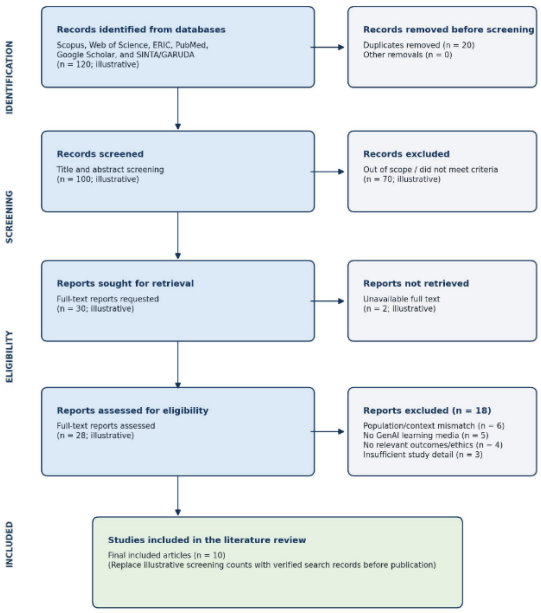


Figure 1. PRISMA 2020 flow diagram

RESULTS AND DISCUSSION

Result

Overview of the Evidence

The evidence matrix below brings together ten empirical studies and research syntheses relevant to generative AI (GenAI) as learning media. It includes experimental and classroom studies, student-perception research, systematic reviews, and a meta-analysis. The studies do not share a single outcome measure or design; therefore, their results are summarized narratively rather than combined into a new pooled estimate. Reported sample sizes and study counts are given where available. When an article did not provide a comparable numeric effect in the accessible report, the result is described qualitatively rather than estimated.

Table 1. Selected evidence on personalization, effectiveness, and ethics

Study	Design and reported evidence base	Main findings relevant to this review
Jauhiainen and Guerra (2024)	Classroom study; 110 pupils, ages 8–14, in grades 4–6 at two schools	ChatGPT-generated texts, illustrations, and exercises were adjusted to pupils' knowledge levels during lessons. The authors reported increased motivation and improved performance, but the study used descriptive analysis and does not provide a pooled effect estimate.
Ng et al. (2024)	Quasi-experimental comparison; 74 Hong Kong Secondary 4 students; three-week science learning activity	Students using the generative AI-supported SRLbot showed gains in science knowledge, behavioural engagement, and motivation. Students valued personalized recommendations and flexibility.
Kinder et al. (2025)	Teacher-education study; 269 German pre-service teachers	Participants rated adaptive ChatGPT feedback as more useful and interesting than static feedback. The study's main result concerns perceptions of feedback; it should not be interpreted on its own as evidence of durable learning gains.
Yilmaz and Yilmaz (2023)	Empirical study of generative AI use in programming education	The experimental group showed significantly higher computational-thinking skills, programming self-efficacy, and lesson motivation than the comparison group. The source summary does not provide sample size or effect estimates.
Essel et al. (2024)	Mixed-methods pretest–posttest control-group study with Ghanaian university students	The study examined critical, creative, and reflective thinking. Its design compared in-class ChatGPT use with conventional databases and search engines; the accessible report describes the measures and design but does not provide comparable numerical effects in the search summary.
Deng et al. (2025)	Systematic review and meta-analysis; 69 experimental studies identified, with 62 studies included in meta-analyses of frequently reported outcomes	ChatGPT interventions were associated with improved academic performance (reported mean effect, g = 0.712) and affective-motivational outcomes (g = 1.122). The authors caution that study quality, assessment conditions, and limited long-term evidence constrain interpretation.
Yusuf et al. (2024)	Systematic mapping review; 407 publications indexed in Web of Science, ScienceDirect, and Scopus	The review mapped a rapidly expanding research landscape across educational levels, disciplines, and application types. It provides field-level context rather than a direct estimate of learning effectiveness.
The rapid rise of generative AI and its implications for academic integrity (2024)	Student-perspective study of GenAI use for assessment assistance	The study centers students' perceptions and use of chatbots in assessed work, highlighting the need to address academic integrity through clear educational responses and policies.
The ethics of using AI in K–12 education (2024)	Systematic literature review	The review identifies ethical concerns in school use of AI, including academic integrity, privacy, and appropriate student and teacher use.
Navigating the ethical terrain of AI in education (2024)	Systematic review of empirical research on responsible, human-centered AI in education	The review organizes ethical considerations around fairness, privacy, non-maleficence, agency, and transparency. These concerns apply to the design and governance of GenAI learning media as well as to conventional AI systems.

Thematic Findings

Across the selected evidence, personalization emerged as a prominent affordance of GenAI. Studies involving school pupils and pre-service teachers suggest that generated explanations, materials, and feedback can be adapted to learners' knowledge or task context. In the science classroom study, students also valued personalized recommendations and flexibility (Jauhiainen & Guerra, 2024; Ng et al., 2024). These findings support personalization as a plausible instructional function, while leaving open how reliably it works across subjects, age groups, and longer courses. Evidence on effectiveness was generally positive, but varied in strength and scope. Individual classroom and programming studies reported improvements in motivation, engagement, or subject-related outcomes. The meta-analysis reported positive average effects for academic performance and affective-motivational outcomes, but also noted methodological limitations and the need for more long-term assessment (Deng et al., 2025). Thus, immediate performance or positive student perceptions should not be treated as equivalent to lasting learning, independent problem-solving, or transfer to new tasks. For ethical challenges, the reviewed literature identifies academic integrity and assessment authenticity alongside privacy, fairness, transparency, and learner agency. This evidence indicates that ethical considerations are not separate from instructional effectiveness: unclear rules or uncritical reliance on generated responses may affect both the credibility of assessment and the development of learner independence. Taken together, the ten studies suggest that GenAI can support personalized learning and show promising educational outcomes, but effectiveness depends on instructional design, critical evaluation, and responsible oversight.

Discussion

The findings suggest that generative artificial intelligence (GenAI) can function as learning media through three related affordances: adapting content, providing responsive assistance, and supporting learners' engagement with tasks. However, these affordances do not automatically produce educational benefits. Their value depends on instructional design, the learner's active role, and safeguards that address reliability and ethical use. This interpretation is consistent with scholarship that treats GenAI as a potential support for teaching and learning while emphasizing its limitations and the need for responsible integration (Kasneci et al., 2023; Tlili et al., 2023; Yusuf et al., 2024).

Personalization as Pedagogical Support

The reviewed studies show how GenAI may personalize learning by varying explanations, materials, and feedback in response to learner needs. Jauhiainen and Guerra (2024), for example, used ChatGPT to adapt learning materials for 110 pupils aged 8–14 across two schools. They reported that generated texts, illustrations, and exercises could be adjusted to pupils' knowledge levels during lessons, alongside positive indications for motivation and performance. In a different context, Ng et al. (2024) compared a generative AI-supported chatbot with a rule-based chatbot among 74 Hong Kong secondary students. Over a three-week science learning activity, students using the generative chatbot demonstrated gains in science knowledge, behavioural engagement, and motivation; students also valued its personalized recommendations and flexibility.

These findings connect personalization with self-regulated learning (SRL). A system that adjusts explanations or prompts may help learners set goals, check understanding, and choose a next step. Yet personalization should not be equated with unrestricted customization or automated decisions about what students need. In the Ng et al. study, the chatbot was embedded in a structured learning design and guided by a teacher. This suggests that GenAI's adaptive responses are most educationally meaningful when they support a deliberate learning process and remain open to teacher review. Reviews of GenAI in education likewise identify differentiated content and individualized support as promising applications, but emphasize that tools must be aligned with curriculum and pedagogy (Wang et al., 2024; Zawacki-Richter et al., 2024).

Feedback is another important part of personalization. In a study of 269 German pre-service teachers, Kinder et al. (2025) found that participants rated adaptive feedback generated by ChatGPT as more useful and interesting than static feedback. This result supports the perceived value of responsive feedback, but perceived usefulness is not itself proof of improved learning. The distinction matters: feedback may motivate learners to revise, but its instructional impact depends on accuracy, clarity, and whether learners understand how to act on it. Consistent with this interpretation, a review of empirical GenAI studies in pedagogical practice identifies feedback and content support among emerging applications while noting the need to evaluate how these tools affect learning in authentic educational settings (Wang et al., 2025).

Effectiveness and the Strength of Evidence

The evidence on effectiveness is promising but uneven. In programming education, Yilmaz and Yilmaz (2023) reported significantly higher computational-thinking skills, programming self-efficacy, and lesson motivation in the group using a generative AI-based tool. Ng et al. (2024) similarly reported improvements in science knowledge, engagement, and motivation in their structured chatbot intervention. These studies suggest that GenAI may support both cognitive outcomes and affective dimensions of learning when learners use it as part of a planned instructional activity. The broader experimental evidence synthesized by Deng et al. (2025) strengthens the case for potential effectiveness. Their review identified 69 experimental studies and meta-analyzed 62 studies for frequently reported outcomes. The authors reported positive average effects for academic performance ($g = 0.712$) and affective-motivational outcomes ($g = 1.122$). These results indicate that ChatGPT-supported interventions can improve measured outcomes in some contexts. They should not, however, be interpreted as a guarantee of equal benefit across age groups, subjects, or intervention designs. Deng et al. also highlight methodological limitations and the need for longer-term evaluation, underscoring the difference between immediate performance and durable learning. This caution is especially important when GenAI supplies answers or completes parts of a task. A learner may produce a more polished response without acquiring the reasoning or subject knowledge that assessment is intended to measure. Therefore, effectiveness should be assessed through more than task completion or learner satisfaction. Measures of retention, transfer, independent problem solving, and students' ability to evaluate generated content are needed to determine whether GenAI is supporting learning or replacing part of the learning process. A growing research landscape, mapped across 407 publications by Yusuf et al. (2024), shows that applications and claims have expanded quickly; the scale of publication activity should not be confused with settled evidence of long-term effectiveness.

Ethical Challenges and Responsible Use

Ethical concerns are integral to the educational value of GenAI. The reviewed literature identifies academic integrity and assessment authenticity, as well as concerns about privacy, bias, transparency, and learner agency. Student perspectives illustrate the complexity of these issues: in one higher education study, students expressed more support for limited writing assistance than for using ChatGPT to write an entire essay (Parsons et al., 2024). This distinction suggests that policy should specify acceptable forms of support rather than treating all AI use as equivalent. Assessment design also matters. When assignments can be completed by generated responses with little learner contribution, the assessment may no longer provide valid evidence of individual understanding. Research on student perceptions and assessment use highlights the need for explicit institutional responses to GenAI-assisted work (Johnston et al., 2024). At the same time, detector-based approaches alone cannot resolve questions about authorship, learning, or appropriate use. More robust practice combines transparent disclosure expectations with assignments that make reasoning, process, and reflection visible. A human-centered approach also requires attention to privacy, fairness, and control. Reviews of responsible AI in education identify fairness and equity, privacy and security, agency and autonomy, and transparency as central concerns (Gouseti et al., 2024; Huo & Siau, 2024). UNESCO (2023) similarly calls for a human-centered approach, including safeguards and capacity-building. Taken together, this evidence suggests that GenAI should support teachers' and learners' decisions without obscuring how outputs are produced or diminishing human responsibility. Students need opportunities to question generated content, recognize possible errors, and make informed choices about what information they share with AI systems.

Integrated Interpretation

Overall, the literature indicates that GenAI can support personalized learning, responsive feedback, and selected academic outcomes, particularly when embedded in structured teaching and learning activities. Empirical studies provide encouraging findings, and meta-analytic evidence reports positive average effects. Nevertheless, the evidence base remains heterogeneous, and perceived usefulness, short-term achievement, and long-term learning are distinct outcomes. Ethical safeguards and sound assessment practices are therefore conditions for meaningful educational use, rather than issues to be considered after adoption. For teachers and institutions, the practical implication is to integrate GenAI purposefully, evaluate its outputs, preserve learner agency, and assess whether students can demonstrate understanding independently. Further empirical research should compare instructional approaches across educational levels, measure retention and transfer, and examine how personalization can be delivered fairly and securely.

CONCLUSION

This literature review indicates that generative artificial intelligence (GenAI) can serve as learning media by adapting explanations, materials, and feedback to learners' needs. Empirical classroom studies provide encouraging evidence: a study of 110 pupils reported that ChatGPT-supported materials could be dynamically tailored to different knowledge levels, while a three-week study involving 74 secondary students found gains in science knowledge, behavioural engagement, and motivation when students used a generative AI-supported self-regulated learning chatbot. A separate study of 269 pre-service teachers found adaptive ChatGPT feedback more useful and interesting than static feedback. At the broader evidence level, a meta-analysis of 69 experimental studies reported positive average effects on academic performance (g = 0.712) and affective-motivational outcomes (g = 1.122).

These findings support GenAI's potential, but the evidence does not establish uniform or lasting benefits across contexts. Its educational value depends on alignment with learning objectives, teacher guidance, reliable outputs, and assessments that reveal learners' own understanding. Ethical concerns—including academic integrity, privacy, fairness, transparency, and learner autonomy—must therefore be addressed as part of implementation. Overall, GenAI is best understood as a supportive learning medium whose benefits depend on responsible pedagogical design. Further research should examine long-term retention, transfer, and equitable use across educational settings.

REFERENCES

Ansari, A. N., Ahmad, S., & Bhutta, S. M. (2024). Mapping the global evidence around the use of ChatGPT in higher education: A systematic scoping review. Education and Information Technologies, 29, 11281–11321. https://doi.org/10.1007/s10639-023-12223-4

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International. https://doi.org/10.1080/14703297.2023.2190148

Decoding AI ethics from users' lens in education: A systematic review. (2024). Heliyon, 10(20), e39357. https://doi.org/10.1016/j.heliyon.2024.e39357

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. Computers & Education, 227, 105224. https://doi.org/10.1016/j.compedu.2024.105224

Essel, H. B., Vlachopoulos, D., Essuman, A. B., & Amankwa, J. O. (2024). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs). Computers and Education: Artificial Intelligence, 6, 100198. https://doi.org/10.1016/j.caeai.2023.100198

Fu, Y., & Weng, Z. (2024). Navigating the ethical terrain of AI in education: A systematic review on framing responsible human-centered AI practices. Computers and Education: Artificial Intelligence, 7, 100306. https://doi.org/10.1016/j.caeai.2024.100306

Gouseti, A., et al. (2024). The ethics of using AI in K–12 education: A systematic literature review. https://doi.org/10.1080/1475939X.2024.2428601

Gouseti, A., James, F., Fallin, L., & Burden, K. (2025). The ethics of using AI in K–12 education: A systematic literature review. Technology, Pedagogy and Education, 34(2), 161–182. https://doi.org/10.1080/1475939X.2024.2428601

Gruenhagen, J. H., Sinclair, P. M., Carroll, J.-A., Baker, P. R. A., Wilson, A., & Demant, D. (2024). The rapid rise of generative AI and its implications for academic integrity: Students' perceptions and use of chatbots for assistance with assessments. Computers and Education: Artificial Intelligence, 7, 100273. https://doi.org/10.1016/j.caeai.2024.100273

Huo, Y., & Siau, K. (2024). Navigating the ethical terrain of AI in education: A systematic review on framing responsible human-centered AI practices. Computers and Education: Artificial Intelligence, 5, 100306. https://doi.org/10.1016/j.caeai.2024.100306

Jauhiainen, J. S., & Guerra, A. G. (2024). Generative AI and education: Dynamic personalization of pupils' school learning material with ChatGPT. Frontiers in Education, 9, 1288723. https://doi.org/10.3389/feduc.2024.1288723

Johnston, H., et al. (2024). The rapid rise of generative AI and its implications for academic integrity: Students' perceptions and use of chatbots for assistance with assessments. Computers and Education: Artificial Intelligence. https://doi.org/10.1016/j.caeai.2024.100273

Kasneci, E., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kinder, A., et al. (2025). Effects of adaptive feedback generated by a large language model: A case study in teacher education. Computers and Education: Artificial Intelligence, 8, 100349. https://doi.org/10.1016/j.caeai.2024.100349

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. Education Sciences, 13(4), 410. https://doi.org/10.3390/educsci13040410

Luo, J. (2024). A critical review of GenAI policies in higher education assessment: A call to reconsider the "originality" of students' work. Assessment & Evaluation in Higher Education, 49, 651–664. https://doi.org/10.1080/02602938.2024.2309963

Montenegro-Rueda, M., Fernández-Cerero, J., Fernández-Batanero, J. M., & López-Meneses, E. (2023). Impact of the implementation of ChatGPT in education: A systematic review. Computers, 12(8), 153. https://doi.org/10.3390/computers12080153

Ng, D. T. K., Tan, C. W., & Leung, J. K. L. (2024). Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study. British Journal of Educational Technology, 55, 1328–1353. https://doi.org/10.1111/bjet.13454

Page, M. J., McKenzie, J. E., 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., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71

Parsons, B., Johnston, H., Boey, T., Shanks, E., & Wells, R. (2024). Student perspectives on the use of generative artificial intelligence technologies in higher education. International Journal for Educational Integrity, 20. https://doi.org/10.1007/s40979-024-00149-4

Short-term learning effect of ChatGPT on pharmacy students' learning. (2024). Research in Social and Administrative Pharmacy, 20. https://doi.org/10.1016/j.rcsop.2024.100478

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learning Environments, 10, 15. https://doi.org/10.1186/s40561-023-00237-x

UNESCO. (2023). Guidance for generative AI in education and research. https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research

Wang, N., Wang, X., & Su, Y. S. (2024). Critical analysis of the technological affordances, challenges and future directions of generative AI in education: A systematic review. Asia Pacific Journal of Education, 44(1), 139–155. https://doi.org/10.1080/02188791.2024.2305156

Wang, X., et al. (2025). Generative artificial intelligence in pedagogical practices: A systematic review of empirical studies (2022–2024). Cogent Education. https://doi.org/10.1080/2331186X.2025.2485499

Yilmaz, R., & Yilmaz, F. G. K. (2023). The effect of generative artificial intelligence (AI)-based-tool use on students' computational thinking skills, programming self-efficacy and motivation. Computers and Education: Artificial Intelligence, 4, 100147. https://doi.org/10.1016/j.caeai.2023.100147

Yusuf, A., Pervin, N., Román-González, M., & Noor, N. M. (2024). Generative AI in education and research: A systematic mapping review. Review of Education, 12, e3489. https://doi.org/10.1002/rev3.3489

Zawacki-Richter, O., Bai, J. Y. H., Lee, K., Slagter van Tryon, P. J., & Prinsloo, P. (2024). New advances in artificial intelligence applications in higher education. International Journal of Educational Technology in Higher Education, 21, Article 32. https://doi.org/10.1186/s41239-024-00464-3